SEMI-ANNUAL STATUS REPORT

Error Control Techniques for Satellite and Space Communications

NASA Grant Number NAG5-557

Principal Investigator:
Daniel J. Costello, Jr.

February 1990

I

Summary of Progress

During the period August 1, 1989 - January 31, 1990, progress was made in the following areas:

1) Performance Analysis of Bandwidth Efficient Trellis Codes

Two methods have traditionally been employed to analyze the performance of various coding schemes. One method bounds the achievable free distance of particular classes of codes, since free distance is the most important parameter that influences the performance of a code. The other method uses a random coding approach to directly bound the average error probability of an ensemble of codes. The best codes are then known to perform at least as well as the bound. This method is the one originally taken by Shannon.

Most of the performance analyses published for trellis coded modulation (TCM) schemes have used the first method, i.e., to bound the achievable free distance of particular classes of codes. We have just completed a new analysis of TCM schemes which uses the random coding approach. A paper summarizing these results has been submitted for publication to the IEEE Transactions on Information Theory [1]. A copy of this paper is included as Appendix A of this report. The most interesting aspect of this paper is that the cutoff rate R_0 of the channel is shown to be the most important factor determining the performance of TCM schemes. This fact can be used to find signal constellations which maximize the performance of a particular class of codes when combined with an appropriate mapping.

We have also continued our work on the performance analysis of concatenation schemes with TCM inner codes and Reed-Solomon (RS) outer codes. Our previous work on this problem, summarized in earlier reports submitted to NASA and detailed in several journal and conference publications, used an approach of simulating the performance of the inner code and then using RS code bounds to determine overall performance. This approach was necessitated by the fact that all previous performance bounds for TCM schemes treated only the bit error probability, whereas for concatenation schemes the symbol error probability of the inner code is the parameter of interest.

We have now developed a new bound on the symbol error probability of trellis codes. A summary of this work, which was recently presented at the 1990 IEEE International Symposium on Information Theory [2], is included as Appendix B of this report. Using this new bound, we are now able to do a complete analysis of TCM/RS concatenation schemes without resorting to simulations. This will allow us to examine the performance of a much greater variety of possible concatenation schemes, since simulation studies are particularly difficult and time consuming for TCM codes. Mr. Lance Perez, a Ph.D. student supported by the grant, is conducting this phase of our research. We plan to submit a paper for publication on this new bound in the near future.

2) Construction of Bandwidth Efficient Trellis Codes

In our annual status report of October 1989, we included the final version of a full length paper in which a large number of new trellis codes were constructed. Most of these codes used multi-dimensional (multi-D) 4-PSK, 8-PSK, and 16-PSK signal constellations, although new codes for two-dimensional (2-D) signal constellations were also given. We have since begun work on the construction of two new classes of trellis codes:

a) Nonlinear 2-D trellis codes which are fully invariant to discrete rotations of the PSK signal set.
b) Multi-D trellis codes for QAM signal sets.

Rotational invariance is a desirable feature for TCM schemes. Rotationally invariant codes have the property that if the demodulator locks onto the wrong phase of the received signal, the decoder will suffer only a slight degradation in performance. (This also assumes the use of differential encoding and decoding.) This is particularly important in applications where the traffic (or the channel) is bursty, thereby causing the demodulator to periodically reacquire phase lock. Unfortunately, no 2-D linear convolutional code can be fully invariant to discrete phase rotations of the signal set. This is one of the motivating factors in considering multi-D signal sets, where it is possible to find linear codes with full rotational invariance. On the other hand, 2-D TCM schemes are much simpler to implement than multi-D schemes and are often required for this reason. This led us to the construction of nonlinear convolutional codes for 2-D signal sets which have full rotational invariance. In general, there is a small price in performance to be paid to guarantee rotational invariance in the 2-D case. A summary of our new nonlinear codes, presented at the 1989 IEEE Workshop on Information Theory [3], is included as Appendix C of this report. This work is being conducted by Mr. Steven Pietrobon, a Ph.D. student supported by the grant. A full length paper is being prepared for submission in the near future which will contain an extensive list of nonlinear rotationally invariant codes for 8-PSK and 16-PSK signal constellations.

In some applications, constant amplitude signals such as PSK may not be required. In this case, other signal constellations such as QAM can be considered. We have extended our constructions of multi-D TCM codes to the QAM case. Generally, better performance can be obtained with QAM than with PSK because there is more flexibility in assigning signal points, thereby making it possible to achieve larger free distances with the same average signal energy. A brief summary of our new QAM code constructions, recently presented at the 1990 IEEE International Symposium on Information Theory [4], is included as Appendix D of this report. This work is being performed by Mr. Steven Pietrobon, a Ph.D. student supported by the grant. A full length paper is being prepared for submission in the near future which will contain extensive lists of multi-D codes for a variety of QAM signal constellations.

3) Sequential Decoding of Trellis Codes

One of the major thrusts of our future research efforts under the grant will be the development of suboptimum decoding methods for TCM schemes. Optimum (Viterbi) decoding can only be used to obtain moderate error rates on the order of $10^{-4} - 10^{-5}$ on many channels. To obtain lower error rates would require the use of prohibitively complex decoders (long constraint or block lengths). Therefore to achieve error rates in the range $10^{-6} - 10^{-9}$ will require the use of longer codes and suboptimum (but still very good) decoding methods which are insensitive to code constraint (block) length. (Another approach to the problem of achieving lower error rates than can be obtained with Viterbi decoding is to use concatenated coding, which is under continuing investigation.)

Sequential decoding has long been recognized as a nearly optimum decoding method whose complexity is insensitive to code constraint length. Therefore sequential decoding can be used with large constraint length codes. One major problem with sequential decoders, however, is that long searches are occasionally necessary, and this may result in some lost or erased data. Therefore, in order to fairly compare sequential decoding with Viterbi decoding, it is necessary to account for the erasures in some way, since Viterbi decoders never erase any information.

We have begun the development of an erasurefree version of sequential decoding which can be directly compared to Viterbi decoding. Some preliminary results of this work, which were presented at the 1990 IEEE International Symposium on Information Theory [5], are included as Appendix E of this report. Our erasurefree sequential decoding algorithm, called the buffer looking algorithm (BLA), appears to perform quite well. Simulation results show that its performance with a constraint length 13, rate 2/3, 8-PSK trellis code is about 1dB superior to Viterbi decoding of a constraint length 8, rate 2/3, 8-PSK trellis code at a decoded error probability of 10^{-5}. At lower error rates, we would expect the relative performance of the sequential decoder to be even better. A complete comparison of the performance, complexity, and delay of sequential decoding and Viterbi decoding of trellis codes will be the subject of future reports, but the preliminary results look very encouraging. Mr. Fu-Quan Wang, a Ph.D. student supported by the grant, is conducting our research on sequential decoding. Dr. Daniel J. Costello, Jr., the principal investigator on the grant, has been asked to give an invited lecture on this research at the 1990 IEEE Information Theory Workshop to be held in Eindhoven, The Netherlands, in June.

References

[1] C. Schlegel and D. J. Costello, Jr., "New Performance Bounds for Trellis Coded Modulation", submitted to *IEEE Transactions on Information Theory,* February 1990.

[2] L. C. Perez and D. J. Costello, Jr., "An Upper Bound on the Symbol Error Rate for Convolutional and Trellis Codes", 1990 IEEE International Symposium on Information Theory, San Diego, CA, January 1990.

[3] S. S. Pietrobon, D. J. Costello, Jr., and G. Ungerboeck, "A General Parity Check Equation for Rotationally Invariant Trellis Codes", 1989 IEEE Information Theory Workshop, Cornell University, Ithaca, NY, June 1989.

[4] S. S. Pietrobon and D. J. Costello, Jr., "Trellis Coding Using Multi-Dimensional QAM Signal Sets", 1990 IEEE International Symposium on Information Theory, San Diego, CA, January 1990.

[5] F. Q. Wang and D. J. Costello, Jr., "Erasurefree Sequential Decoding and Its Application to Trellis Codes", 1990 IEEE International Symposium on Information Theory, San Diego, CA, January 1990.

Appendix A

New Performance Bounds for Trellis Coded Modulation

New Performance Bounds for Trellis Coded Modulation*

Christian Schlegel†
Communications Group
ASEA BROWN BOVERI Corporate Research
5405 Baden Switzerland
and
Daniel J. Costello, Jr.
Department of Electrical and Computer Engineering
University of Notre Dame
Notre Dame, IN 46556, USA

February 1, 1990

Abstract

This paper presents an expurgated upper bound on the event error probability of trellis coded modulation. This bound is used to derive a lower bound on the minimum achievable free Euclidean distance d_{free} of trellis codes. It is shown that the dominant parameters for both bounds, the expurgated error exponent and the asymptotic d_{free} growth rate, respectively, can be obtained from the cutoff-rate R_0 of the transmission channel by a simple geometric construction, making R_0 the central parameter for finding good trellis codes. Several constellations are optimized with respect to the bounds.

*This work was supported by NASA Grant NAG5-557 and NSF Grant NCR89-03429.

†Formerly with the Department of Electrical and Computer Engineering, University of Notre Dame, Notre Dame, IN 46556.

I. Introduction

In recent years bandwidth efficient *trellis coded modulation* (TCM) has become increasingly popular and much analysis has been devoted to the performance of these coding schemes on AWGN-channels (see [1-5] and the references therein). It is well known that for large signal-to-noise ratio (SNR), the minimum free Euclidean distance d_{free} of a trellis code is the dominant parameter of a code's performance. Much research has gone into the search for and the construction of codes with large d_{free}. While most of this work has focused on finding good trellis codes with a given signal constellation, the constellation itself is also a parameter in the system design. There have been a few attempts to design codes using non-standard signal constellations, like the asymmetric MPSK signal sets introduced in [6]. These codes showed slight performance improvements, but no general rule on how to choose a constellation is known.

In this paper we show that a signal constellation with a good value of the cutoff-rate R_0 [7] will indicate the existence of codes with good d_{free} and good performance. This is done by calculating an expurgated upper bound on the first event error probability of a trellis code and relating it to d_{free}.

A code's minimum free Euclidean distance d_{free}[1] is often used to obtain an estimate of the code's error performance as follows:

$$P_e \approx n_{free} \mathrm{Q}\left(d_{free} \sqrt{\frac{E_s/N_0}{2}} \right),$$

where n_{free} is the path multiplicity of the code, i.e., the number of error events with distance d_{free}, and $Q(x) = \int_x^\infty 1/\sqrt{2\pi}\exp(x^2/2)dx$. This approximation provides a good asymptotic estimate of a code's performance.

This paper is organized in the following way. Section II describes TCM and the definitions used later. In Sections III and IV we derive a random coding bound and an expurgated bound on the first event error probability of TCM. The casual reader may want to skip this derivation and proceed directly to Theorem 1 in Section IV. In Section V we present a strict lower bound on the event error probability involving d_{free}, and, relating it to the expurgated upper bound, we rederive the lower bound on d_{free} originally presented by Rouanne and Costello [8]. In Section VI we develop a geometric approach to constructing the bounds and determine a number of optimized constellations. Section VII contains the conclusions.

II. Trellis Coded Modulation

A general TCM communication system (Figure 1) consists of a trellis encoder, a modulator, the transmission channel, a demodulator, and a trellis decoder. The structure of a trellis code is generated by a binary convolutional encoder, which is a finite state automaton with 2^ν possible states, where ν is the *total memory* of the encoder. In the minimal realization [9], the encoder consists of $\tilde{k}$ feedback free shift register chains of lengths $\nu_1, \ldots, \nu_{\tilde{k}}$. We

[1]Note that all Euclidean distances are normalized, i.e., they are based on unit energy signal constellations.

assume in this papers that $\nu_1 = \nu_2 = \cdots = \nu_{\tilde{k}} = \nu_m$, where ν_m is the *memory length* of the code. It then follows that the shortest non-zero path has length $\mu = \nu_m + 1$. μ is called the *constraint length* of the code. An extension to different values of ν_i is generally possible but messy, and does not seem to provide any additional insight. At each time interval r, the encoder accepts $\tilde{k}$ binary input bits $(u_4^{\tilde{k}}, u_r^{\tilde{k}-1}, \ldots, u_r^1)$ and makes a transition from its state S_r at time r to one of $2^{\tilde{k}}$ possible successor states S_{r+1} at time $r+1$.

The $\tilde{n} = n - (k - \tilde{k})$ output bits from the convolutional encoder and $k - \tilde{k}$ uncoded information bits $(u_r^k, \ldots, u_r^{\tilde{k}+1})$ form one of 2^n binary n-tuples $v_r = (v_r^n, v_r^{n-1}, \ldots, v_r^1)$, called a signal selector. The sequence $V = (v_1, \ldots, v_l)$ of signal selectors is *the label* of a path through a linear trellis[2], generated by the convolutional encoder. v_r is then mapped into x_r, one of $A = 2^n$ d-dimensional channel symbols from a signal set $\mathcal{A} = \{a_1, a_2, \ldots, a_A\}$ of cardinality A.

The uncoded information bits do not affect the state of the convolutional encoder and cause $2^{k-\tilde{k}}$ parallel transitions between the encoder states S_r and S_{r+1}. A rate $R = k/n$ trellis code transmits k bits/channel signal.

In practical systems, one often uses 2-dimensional (complex) signal sets for their ease of implementation, and the real part and imaginary part of x_r drive the direct and quadrature component of the modulator.

III. A Random Coding Bound for Time Varying Trellis Codes on General Memoryless Channels

In this section we derive an expurgated upper bound on the event error probability of a trellis code. The derivation is similar to that given in Viterbi and Omura [10] for convolutional codes. Throughout the derivation we assume that the codes are used in conjunction with a maximum-likelihood decoder that operates on a decoding metric $m(\mathbf{x}, \mathbf{y})$, where $\mathbf{x} = (x_1, \ldots, x_l)$ is a sequence of transmitted symbols x_i and $\mathbf{y} = (y_1, \ldots, y_l)$ is the corresponding received symbol sequence. By convention, the signal $\mathbf{x}$ with the lowest metric is the most reliable, i.e., $m(\mathbf{x}, \mathbf{y})$ is some non-negative function of $\mathbf{x}$ given $\mathbf{y}$, which is inversely related to the conditional probability that $\mathbf{x}$ was transmitted given that $\mathbf{y}$ was received. The decoder then chooses the message sequence $\mathbf{x}$ for which this metric is minimized. It makes an error if it decodes a sequence $\mathbf{x}'$, given that the correct sequence, i.e., the transmitted sequence, was $\mathbf{x}$. This happens if $m(\mathbf{x}', \mathbf{y}) \le m(\mathbf{x}, \mathbf{y})$.

Let V and V' be labeled paths through the trellis, i.e., V and V' describe trellis paths without signals assigned to them. We refer to V as the correct path if it is the one followed by the encoder. Let V' be a path that diverges from V at node j. We call V' an incorrect path. Further, let $\mathcal{V}'$ be the set of all incorrect paths V' that diverge from V at node j. The paths V' eventually remerge with V and we call the number of branches over which V and V' differ the length of V'. Due to the linearity of the labeling, the sets $\mathcal{V}'$ for different correct paths V are equivalent, i.e., they contain the same number of paths of the same lengths. In

[2]Here *linear* means that if the binary output sequence V of the convolutional encoder is used to label a path in the trellis, the modulo-2 sum of two labels is a label for another valid path.

a particular trellis code, let $\mathbf{x}$ be the sequence of signals assigned to the correct path V, and let $\mathbf{x}'$ be the sequence of signals assigned to V'.

Our goal is to obtain an upper bound on the *first event error probability* $P_e(j)$, the probability that the decoder starts an error event at node j. An error event starts at node j if the decoder chooses an incorrect path V' with its associated signal sequence $\mathbf{x}'$ over the correct path V with signal sequence $\mathbf{x}$ starting at node j, as illustrated in Figure 2.

A necessary but not sufficient condition for such an error event to occur is that the incorrect path V' accumulates a smaller total metric than the correct path V over their unmerged segments or time intervals of the trellis. The probability $P_e(j)$ may then be upper bounded by the probability that any path $V' \in \mathcal{V}'$ diverging from the correct path V at node j accumulates a lower total metric than the correct path V. This probability must then be averaged over all correct paths V. Letting $p(V)$ denote the probability of path V, we obtain

$$P_e(j) \leq \sum_V p(V) \sum_{\mathbf{y}} p(\mathbf{y}|\mathbf{x}) \mathcal{I} \left\{ \bigcup_{V' \in \mathcal{V}'} V'(m(\mathbf{x}', \mathbf{y}) - m(\mathbf{x}, \mathbf{y}) \leq 0) \right\}, \tag{1}$$

where $V'(m(\mathbf{x}', \mathbf{y}) - m(\mathbf{x}, \mathbf{y}) \leq 0)$ is a path $\in \mathcal{V}'$ for which $m(\mathbf{x}', \mathbf{y}) - m(\mathbf{x}, \mathbf{y}) \leq 0$, and $\mathcal{I}(B)$ is a set indicator function such that $\mathcal{I}(B) = 0$ if $B = \emptyset$, the empty set, and $\mathcal{I}(B) = 1$ if $B \neq \emptyset$. $p(\mathbf{y}|\mathbf{x})$ is the conditional probability of receiving sequence $\mathbf{y}$ if the encoder follows path V and transmits the signal sequence $\mathbf{x}$. This conditional probability depends on the particular channel over which the sequences are transmitted.

If the received signal sequence $\mathbf{y}$ consists of real valued symbols, rather than discrete signal points (unquantized decoding), the summation in (1) is replaced by an integration over the space of $\mathbf{y}$, i.e.,

$$P_e(j) \leq \sum_V p(V) \int_{\mathbf{y}} p(\mathbf{y}|\mathbf{x}) \mathcal{I} \left\{ \bigcup_{V' \in \mathcal{V}'} V'(m(\mathbf{x}', \mathbf{y}) - m(\mathbf{x}, \mathbf{y}) \leq 0) \right\} d\mathbf{y}. \tag{2}$$

It is, in general, too difficult to evaluate (1) or (2) exactly and we therefore resort further bounding techniques. Using the inequality $\mathcal{I}\{\bigcup_i B_i\} \leq \sum_i \mathcal{I}\{B_i\}$, we may immediately simplify (2) to obtain an upper bound of the form:

$$P_e(j) \leq \sum_V p(V) \int_{\mathbf{y}} p(\mathbf{y} \mid \mathbf{x}) \sum_{V' \in \mathcal{V}'} \mathcal{I}\{V'(m(\mathbf{x}', \mathbf{y}) - m(\mathbf{x}, \mathbf{y}) \leq 0)\} d\mathbf{y}. \tag{3}$$

In order for an incorrect path V' to merge with the correct path V at node $j + l$, the last ν_m entries in the information sequences $\mathbf{u}'^1, \ldots, \mathbf{u}'^k$ associated with V' must equal the last ν_m entries in the information sequences $\mathbf{u}^1, \ldots, \mathbf{u}^k$ associated with V, i.e., $u_r'^i = u_r^i$ for $r \in \{j + l - \nu_m, \ldots, j + l - 1\}$ and $i = 1, 2, \cdots, k$. That this is the case can easily be seen by noting that in order for the two paths V' and V to merge at node $j + l$, their associated encoder states must be identical. Because an information bit entering the encoder can affect the output for ν_m time units, this is also the time it takes to force the encoder into any given state from any arbitrary starting state; in particular, to have V' join V at node $j + l$. Because the remaining information bits $u_r'^i$ for $r \in \{j, \ldots, j + l - \mu\}$ are arbitrary, we have

$M \le (2^k - 1)2^{k(l-\mu)}$ incorrect paths V' of length l. (Note that the choice of the information bits at $r = j$ is restricted because we stipulated that the incorrect path diverges at node j, which rules out the one path that continues to the correct state at node $j+1$. This accounts for the term $2^k - 1$ in the expression for M.)

We now proceed to evaluate $\int_{\mathbf{y}} p(\mathbf{y}|\mathbf{x})\mathcal{I}\{V'(m(\mathbf{x}',\mathbf{y}) - m(\mathbf{x},\mathbf{y}) \le 0)\}$ for a particular path pair (V', V) of length l. Let us write (3) as

$$P_e(j) \le \sum_V p(V) \sum_{V' \in \mathcal{V}'} \Pr(\mathbf{x} \to \mathbf{x}'), \tag{4}$$

where

$$\begin{aligned} P(\mathbf{x} \to \mathbf{x}') &\triangleq \int_{\mathbf{y}} p(\mathbf{y}|\mathbf{x})\, \mathcal{I}\{V'(m(\mathbf{x}',\mathbf{y}) - m(\mathbf{x},\mathbf{y}) \le 0)\} dy \\ &= E_{\mathbf{y}|\mathbf{x}}\left[\mathcal{I}\{V'(m(\mathbf{x}',\mathbf{y}) - m(\mathbf{x},\mathbf{y}) \le 0)\}\right], \end{aligned}$$

and $E_{\mathbf{y}|\mathbf{x}}$ denotes conditional expectation. We now use the Chernoff bounding technique [11] and overbound $\mathcal{I}[\alpha \le 0]$ by $\exp(-\lambda\alpha)$ to obtain

$$P(\mathbf{x} \to \mathbf{x}') \le E_{\mathbf{y}|\mathbf{x}}\left[\exp(-\lambda\{m(\mathbf{x}',\mathbf{y}) - m(\mathbf{x},\mathbf{y})\})\right] \triangleq C(\mathbf{x},\mathbf{x}',\lambda),$$

where λ is a non-negative real valued parameter over which $C(\mathbf{x},\mathbf{x}',\lambda)$ is minimized to obtain the tightest possible bound. We call $C(\mathbf{x},\mathbf{x}',\lambda)$ the Chernoff bound between the signal sequences $\mathbf{x}'$ and $\mathbf{x}$.

We now express (4) as the sum over individual sequences of length l

$$\begin{aligned} P_e(j) &\le \sum_V p(V) \sum_{l=\mu}^{\infty} \sum_{V_l' \in \mathcal{V}_l'} C(\mathbf{x},\mathbf{x}',\lambda) \\ &= \sum_{l=\mu}^{\infty} \sum_{V_l \in \mathcal{V}_l} p(V_l) \sum_{V_l' \in \mathcal{V}_l'} C(\mathbf{x},\mathbf{x}',\lambda), \end{aligned} \tag{5}$$

where $\mathcal{V}_l$ is the set of all correct paths V_l of length l starting at node j and $\mathcal{V}_l'$ is the set of all incorrect paths V_l' of length l unmerged with V_l from node j to node $j+l$. Note that $\bigcup_l \mathcal{V}_l' = \mathcal{V}'$.

$P_e(j)$ is the event error probability of a particular code since it depends on the signal sequences $\mathbf{x}$ and $\mathbf{x}'$ of the code. The aim of this section is to obtain a bound on an ensemble of trellis codes, and we therefore must average over the event error probabilities of all the codes in the ensemble, i.e.,

$$\overline{P_e(j)} \le \overline{\sum_{l=\mu}^{\infty} \sum_{V_l \in \mathcal{V}_l} p(V_l) \sum_{V_l' \in \mathcal{V}_l'} C(\mathbf{x},\mathbf{x}',\lambda)}, \tag{6}$$

where the overbar denotes an ensemble average.

Using the linearity of the expectation operator and noting that there are exactly $N = 2^{kl}$ equiprobable paths in $\mathcal{V}_l$, because at each time interval there are 2^k possible choices to continue the correct path, we obtain

$$\begin{aligned}\overline{P_e(j)} &\leq \sum_{l=\mu}^{\infty} \frac{1}{2^{kl}} \overline{\sum_{V_l \in \mathcal{V}_l} \sum_{V_l' \in \mathcal{V}_l'} C(\mathbf{x}, \mathbf{x}', \lambda)} \\ &= \sum_{l=\mu}^{\infty} \overline{\pi_l(j)},\end{aligned} \tag{7}$$

where we have implicitly defined $\overline{\pi_l(j)}$.

We will now proceed to evaluate $\overline{\pi_l(j)}$. Let $\mathbf{x}_1, \cdots, \mathbf{x}_N$ be a set of possible correct signal sequences associated with the paths $V_l \in \mathcal{V}_l$ as we go through the codes in the ensemble and let $q_{lN}(\mathbf{x}_1, \cdots, \mathbf{x}_N)$ be their probability of occurrence. Note that there are M incorrect paths $V_l' \in \mathcal{V}_l'$ with signal sequences $\mathbf{x}_1', \cdots, \mathbf{x}_M'$ that spread around each correct path V_l. Because each incorrect path in $\mathcal{V}_l'$ is also a possible correct path V_l of length l, we have $\mathcal{V}_l' \subset \mathcal{V}_l$. Averaging over all codes in the ensemble is the same as averaging over all possible signal sequences in these codes, i.e., over all assignments of signal sequences $\mathbf{x}$ to paths V. We then obtain

$$\begin{aligned}\overline{\pi_l(j)} &\leq \frac{1}{2^{kl}} \sum_{\mathbf{x}_1} \cdots \sum_{\mathbf{x}_N} q_{lN}(\mathbf{x}_1, \cdots, \mathbf{x}_N) \sum_{h=1}^{N} \sum_{i=1}^{M} C(\mathbf{x}_h, \mathbf{x}_i', \lambda) \\ &= \frac{1}{2^{kl}} \sum_{h=1}^{N} \sum_{i=1}^{M} \sum_{\mathbf{x}_h} \sum_{\mathbf{x}_i'} q_{l2}(\mathbf{x}_h, \mathbf{x}_i') C(\mathbf{x}_h, \mathbf{x}_i', \lambda),\end{aligned} \tag{8}$$

where in the last step we have summed over all pairs of sequences $\mathbf{x}, \mathbf{x}' \neq \mathbf{x}_h, \mathbf{x}_i'$. We have now obtained a bound where we can limit our attention to one correct signal sequence $\mathbf{x}_h$ and one incorrect signal sequence $\mathbf{x}_i'$, both of length l.

In order to proceed further, we will now restrict our attention to memoryless channels. On a memoryless channel, the metrics become additive over the individual time units, i.e.,

$$m(\mathbf{x}, \mathbf{y}) = \sum_{r=1}^{l} m(x_r, y_r).$$

This allows us to rewrite (5) as

$$C(\mathbf{x}, \mathbf{x}', \lambda) = \prod_{r=1}^{l} C(x_r, x_r', \lambda) = \prod_{r=1}^{l} \quad E_{y_r|x_r}\left[\exp(-\lambda\{m(x_r', y_r) - m(x_r, y_r)\})\right],$$

where $C(x_r, x_r', \lambda)$ is the *Chernoff factor* between the signals x_r' and x_r.

We now assume further that in composing our code ensemble, each individual signal in each sequence is chosen independently according to a common probability distribution $q(x)$, i.e., $q_l(\mathbf{x}_h) = \prod_{r=1}^{l} q(x_{hr})$ and $q_l(\mathbf{x}'_i|\mathbf{x}_h) = \prod_{r=1}^{l} q(x_{ir}')$, respectively. In order to make this possible we must assume that the trellis codes are time-varying in nature, for otherwise each symbol would also depend on the choices of the ν_m last symbols. We now obtain a much simpler version of the above bound, namely

$$\overline{\pi_l(j)} \leq \frac{1}{2^{kl}} \sum_{h=1}^{N} \sum_{i=1}^{M} \sum_{\mathbf{x}} \sum_{\mathbf{x}'} \prod_{r=1}^{l} q(x_r) q(x_r') C(x_r, x_r', \lambda). \tag{9}$$

Because the choice of the signals x_r and x'_r does not depend on the particular signal sequences $\mathbf{x}_h$ and $\mathbf{x}'_i$, we dropped the dependency on h and i in (9). Upon interchanging multiplication and summation we obtain

$$\overline{\pi_l(j)} \leq \frac{1}{2^{kl}} \sum_{h=1}^{N} \sum_{i=1}^{M} \prod_{r=1}^{l} \sum_{x_r} \sum_{x'_r} q(x_r) q(x'_r) C(x_r, x'_r, \lambda). \tag{10}$$

The signals x_r, x'_r are chosen randomly from the signal set $\mathcal{A} = \{a_1, \ldots, a_A\}$, where $p(a_p)$ is the probability of choosing a_p, i.e., $q(x_r) = p(a_p)$ if $x_r = a_p$. We may now rewrite (10) as

$$\begin{aligned}
\overline{\pi_l(j)} &\leq \frac{1}{2^{kl}} \sum_{h=1}^{N} \sum_{i=1}^{M} \prod_{r=1}^{l} \sum_{m=1}^{A} \sum_{p=1}^{A} p(a_m) p(a_p) C(a_m, a_p, \lambda) \\
&= \frac{1}{2^{kl}} \sum_{h=1}^{N} \sum_{i=1}^{M} \left(\sum_{m=1}^{A} \sum_{p=1}^{A} p(a_m) p(a_p) C(a_m, a_p, \lambda) \right)^l \\
&\leq (2^k - 1) 2^{k(l-\mu)} \left(\sum_{m=1}^{A} \sum_{p=1}^{A} p(a_m) p(a_p) C(a_m, a_p, \lambda) \right)^l .
\end{aligned}$$

Let us now define $R_0(p)$ as

$$R_0(p) \triangleq -\log_2 \min_{\lambda} \sum_{m=1}^{A} \sum_{p=1}^{A} p(a_m) p(a_p) C(a_m, a_p, \lambda). \tag{11}$$

We may now finally evaluate the average event error probability $\overline{P_e(j)}$ at time unit j as

$$\begin{aligned}
\overline{P_e(j)} &\leq \textstyle\sum_{l=\mu}^{\infty} \overline{\pi_l(j)} \\
&\leq (2^k - 1) 2^{-\mu R_0(p)} \sum_{s=0}^{\infty} 2^{ks} 2^{-sR_0(p)} \\
&= \frac{(2^k - 1) 2^{-\mu R_0(p)}}{1 - 2^{-(R_0(p) - k)}} \quad ; \quad 0 < k < R_0(p).
\end{aligned}$$

Since k is the number of information bits transmitted in one channel symbol x_r, we may call it the information rate in bits per channel use and denote it by the symbol R. $P_e(j)$ is independent of the node j and we may thus drop the parameter j and obtain

$$\overline{P_e} \leq \frac{(2^R - 1) 2^{-\mu R_0(p)}}{1 - 2^{-(R_0(p) - R)}} \quad ; \quad 0 < R < R_0(p). \tag{12}$$

The parameter

$$R_0 \triangleq \max_p R_0(p) = \max_p \left(-\log_2 \min_{\lambda} \sum_{m=1}^{A} \sum_{p=1}^{A} p(a_m) p(a_p) C(a_m, a_p, \lambda) \right) \tag{13}$$

is the *cutoff-rate* of the channel and (12) holds for all rates $R < R_0(p)$. We will later use the uniform distribution $p = 1/A$ in (13) and refer to $R_0 = R_0(1/A)$ as the cutoff rate[3] unless noted otherwise, even though the strict definition of cutoff-rate is (13).

Note that R_0 depends on the particular metric $m(y_r, x_r)$ which is used by the decoder. If the decoder uses the *maximum-likelihood* (ML) metric for a memoryless channel, i.e.,

$$\begin{aligned} m(\mathbf{y}, \mathbf{x}) &= -\log(\Pr(\mathbf{y}|\mathbf{x})) = -\log \prod_{r=1}^{l} \Pr(y_r|x_r) \\ &= \sum_{r=1}^{l} (-\log(\Pr(y_r|x_r))) = \sum_{r=1}^{l} m(x_r, y_r), \end{aligned}$$

(13) becomes the channel cutoff-rate for the optimum receiver, which is the usual definition of R_0 [7]. We will denote the value of λ which maximizes the (13) by λ_{R_0}. In this case, the Chernoff factors will be written as $C(a_m, a_p) \triangleq C(a_m, a_p, \lambda_{R_0})$.

The actual evaluation of the maximum-likelihood metric for most channels is not simple, however. In fact, only for the AWGN-channel does the maximum-likelihood metric assume a form simple enough to be implemented in decoding circuits [7]. For the AWGN-channel, the maximum-likelihood metric is the squared Euclidean distance between the received sequence $\mathbf{y}$ and the transmitted sequence $\mathbf{x}$, i.e., $m(\mathbf{y}, \mathbf{x}) = \sum_{r=1}^{l} |y_r - x_r|^2$. With this metric (13) is minimized by setting $\lambda = \lambda_{R_0} = 1/(2N_0)$, and the Chernoff factors turn out to be exponentials in the squared Euclidean distance, i.e.,

$$C(x_r, x'_r) = e^{-\frac{E_S}{4N_0}|x_r - x'_r|^2},$$

where E_S is the average signal energy.

From this it is easily seen that a code's performance is dominated by the two distinct sequences $\mathbf{x}_1$ and $\mathbf{x}_2$ that are closest to each other in terms of squared Euclidean distance. Their distance is referred to as the *minimum free squared Euclidean distance*, or d^2_{free}, of the code, defined as

$$d^2_{free} \triangleq \min_{\mathbf{x}_1 \neq \mathbf{x}_2} \sum_{r=1}^{l} |x_{1r} - x_{2r}|^2.$$

Figure 3 shows R_0 for the AWGN-channel as a function of the ratio of the average signal energy E_S over the average noise power N_0 for a number of popular signal constellations. It is interesting to note that rectangular constellations fare slightly better than constant envelope constellations with the same number of signal points. The reason for this lies in the added flexibility provided by the amplitude modulation in the case of rectangular constellations.

IV. Expurgated Error Bound

In this section we derive an expurgated bound on the event error probability which improves the R_0-bound, especially for rates R significantly below R_0. The event error probability for

[3] $R_0(1/A)$ is sometimes referred to as the *symmetric cutoff-rate*.

a particular correct path V_c is a special case of (4), i.e.,

$$P_{e|V_c}(j) \leq \sum_{V' \in \mathcal{V}'} C(\mathbf{x}, \mathbf{x}', \lambda). \tag{14}$$

Applying the inequality (see e.g. [11])

$$\sum_i a_i \leq \left(\sum_i a_i^s \right)^{1/s}, \quad 0 < s \leq 1,$$

to (14) we obtain

$$P_{e|V_c}^s(j) \leq \sum_{V' \in \mathcal{V}'} C^s(\mathbf{x}, \mathbf{x}', \lambda).$$

Following analogous steps as those leading from (5) to (7), we obtain

$$\begin{aligned} \overline{P_{e|V_c}^s(j)} &\leq \sum_{l=\mu}^{\infty} \overline{\sum_{V_l' \in \mathcal{V}_l'} C^s(\mathbf{x}, \mathbf{x}', \lambda)} \\ &= \sum_{l=\mu}^{\infty} \overline{\pi_l(j, s, V_c)}, \end{aligned}$$

where for memoryless channels $\overline{\pi_l(j, s, V_c)}$ is given by

$$\begin{aligned} \overline{\pi_l(j, s, V_c)} &\leq \sum_{i=1}^{M} \sum_{\mathbf{x}_h} \sum_{\mathbf{x}_i'} q_{l2}(\mathbf{x}_h, \mathbf{x}_i') C^s(\mathbf{x}_h, \mathbf{x}_i', \lambda) \\ &\leq (2^k - 1) 2^{k(l-\mu)} \left(\sum_{m=1}^{A} \sum_{p=1}^{A} p(a_m) p(a_p) C^s(a_m, a_p, \lambda) \right)^l . \end{aligned}$$

Note that $\overline{\pi_l(j, s, V_c)}$ is independent of V, i.e.,

$$\overline{\pi_l(j, s, V_c)} = \overline{\pi_l(j, s)} \quad ; \quad \text{for all } V_c \in \mathcal{V}$$

and

$$\overline{P_{e|V_c}^s(j)} = \overline{P_e^s(j)},$$

i.e., $P_e^s(j)$ averaged over all time-varying trellis codes is independent of the correct path through the trellis and averaging over all correct paths becomes trivial. We now define $E(s)$ as

$$E(s) \triangleq -\log_2 \min_{\lambda} \sum_{m=1}^{A} \sum_{p=1}^{A} p(a_m) p(a_p) C^s(a_m, a_p, \lambda),$$

and proceed to obtain

$$\begin{aligned} \overline{P_e^s(j)} &\leq \sum_{l=\mu}^{\infty} \overline{\pi_l(j, s)} \\ &< (2^k - 1) 2^{-\mu E(s)} \sum_{t=0}^{\infty} 2^{kt} 2^{-tE(s)} \\ &= \frac{(2^k - 1) 2^{-\mu E(s)}}{1 - 2^{-(E(s)-k)}} \quad ; \quad 0 < k < E(s). \end{aligned}$$

$\overline{P_e^s(j)}$ is the the event error probability $P_e(j)$ raised to the power s, averaged over all codes and all correct sequences. There must then be at least one code in the ensemble for which $P_e^s(j) \leq \overline{P_e^s(j)}$. Using this in the equation above we obtain an expurgated upper bound on the event error probability of the *best* trellis code in the ensemble

$$P_e \leq \overline{P_e^s}^{1/s} \leq \left(\frac{(2^k - 1)}{1 - 2^{-(E(s)-k)}} \right)^{1/s} 2^{-\mu \frac{E(s)}{s}} \qquad ; \quad 0 < k < E(s) 0 < s \leq 1 ,$$

where we have again dropped the dummy parameter j. It is sometimes convenient to express this bound as a function of the memory order ν_m of a code. Since $\nu_m = \mu - 1$, we obtain

Theorem 1: *There exists a rate $R = k$ trellis code, with a trellis generated by a convolutional encoder with register lengths $\nu_i = \nu_m$ for $1 \leq i \leq k$, using a signal constellation $\mathcal{A} = \{a_0, \cdots, a_{A-1}\}$ of cardinality A, whose error event probability P_e is bounded above by*

$$P_e \leq \left(\frac{(1 - 2^{-R})}{2^{E(s)-R} - 1} \right)^{1/s} 2^{-\nu_m \frac{E(s)}{s}},$$

for any s and R such that $0 < s \leq 1$ and $0 < R < E(s)$, where

$$E(s) \triangleq -\log_2 \min_{\lambda} \sum_{m=1}^{A} \sum_{p=1}^{A} p(a_m) p(a_p) C^s(a_m, a_p, \lambda).$$

V. Bounds involving the Free Euclidean Distance

In this section we restrict ourselves to AWGN-channels, the most widely used channel model. All results, however, can be extended to general memoryless channels. The following theorem gives a strict lower bound on the average first error event probability P_e, i.e., $P_e(j)$ averaged over all time units j, for trellis codes used on an AWGN-channel.

Theorem 2: *The average event error probability P_e of a trellis code on an AWGN-channel with one sided noise power spectral density N_0 is lower bounded by*

$$P_e \geq \frac{1}{l_{\max}} \sum_{i=1}^{n_d} p_i Q\left(d_i \sqrt{\frac{E_S}{2N_0}} \right),$$

where d_i is the minimum normalized Euclidean distance $d_i = \min_{\mathbf{x},\mathbf{x}'} \sqrt{\frac{\sum_{r=1}^{l_i} |x_r - x_r'|^2}{E_S}}$ achievable between a particular correct sequence $\mathbf{x}$ and any incorrect sequence $\mathbf{x}'$, $E_S = \sum_{i=1}^{A} p(a_i)|a_i|^2$ is the average signal energy, $l_{\max} = \max_i l_i$, where l_i is the minimum length (in branches) of the error events that achieve the minimum distance d_i, and p_i is the probability that the minimum distance sequence pair $\mathbf{x}, \mathbf{x}'$ has distance d_i. n_d is the number of different minimum distances d_i achievable in a particular code, where $d_1 < d_2 < \cdots < d_{n_d}$, and $d_1 = d_{free}$ is the minimum free Euclidean distance of the code.

Remark: Since trellis codes, in general, are non-linear [1-3], the minimum Euclidean distance d_i among all error paths V' with respect to a particular correct path V depends on V. Then p_i is the fraction of correct paths whose nearest error path is at distance d_i.

Proof: Assume that we want to bound the event error probability at node j (Figure 4). Let $\mathcal{V}'(j)$ be the set of paths diverging from the correct path at node j. Let $\mathbf{x}'$ be the signal sequence on a path $V_l' \in \mathcal{V}_l'(j)$ of length l. Assume that the correct path is V. Denote the probability that the decoder follows V for at least l time units by P_c. Further, denote the set of error paths diverging from V at node $j+r$ by $\mathcal{V}'(j+r)$, and let P_ε be the probability that the decoder chooses any path in the set $\mathcal{E} \triangleq \mathcal{V}'(j) \cup \mathcal{V}'(j+1) \cup \cdots \cup \mathcal{V}'(j+l-1)$, i.e., P_ε is the probability that the decoder diverges from the correct path before node $j+l$. P_ε is lower bounded by

$$P_\varepsilon \geq P(\mathbf{x} \to \mathbf{x}'). \tag{15}$$

This follows from the fact that eliminating all signal sequences but $\mathbf{x}'$ from $\mathcal{E}$ allows us to expand the decision regions of both $\mathbf{x}$ and $\mathbf{x}'$, thus increasing P_c and decreasing $1 - P_c = P_\varepsilon$.

On the other hand, P_ε may be upper bounded tightly by

$$P_\varepsilon \leq P_e(j) + P_e(j+1) + \cdots + P_e(j+l-1). \tag{16}$$

In order to proceed further, we combine (15) and (16) and average over all possible time units and correct paths V[4], i.e.,

$$\overline{P(\mathbf{x} \to \mathbf{x}')} \leq \overline{P_e(j) + P_e(j+1) + \cdots + P_e(j+l-1)}.$$

Due to the linearity of the expectation operator

$$\overline{P_e(j) + P_e(j+1) + \cdots + P_e(j+l-1)} = \overline{P_e(j)} + \overline{P_e(j+1)} + \cdots + \overline{P_e(j+l-1)} = l\overline{P_e(j)},$$

since the average first event error probability $\overline{P_e(j+r)}$ is independent of time when averaged over all possible time units and correct paths. If we denote this average first event error probability by P_e we obtain from above

$$\overline{P(\mathbf{x} \to \mathbf{x}')} \leq lP_e. \tag{17}$$

Note that (17) holds for any incorrect path $\in \mathcal{V}'$ and that l is the length of this path.

We now also carry out the averaging on the left hand side of (17), where in each case we choose the incorrect sequence $\mathbf{x}'$ such that $|\mathbf{x} - \mathbf{x}'|$ is minimized, which yields the tightest possible lower bound. This sequence has length l_i which possibly differs from l in (17). This causes the dilemma that the chosen error sequences $\mathbf{x}'$ may not all have equal lengths l_i, raising the question of which l to use in (17). To guarantee that the bound in (17) is not violated, we let l be the maximum length of the incorrect paths chosen, denoted by $l_{\mathbf{max}}$. For the AWGN-channel with one-sided noise power spectral density N_0, the two code word error probability $P_2(\mathbf{x} \to \mathbf{x}')$ is given by

$$P_2(\mathbf{x} \to \mathbf{x}') = Q\left(\frac{|\mathbf{x} - \mathbf{x}'|}{\sqrt{2N_0}}\right),$$

[4]Here the overbar denotes the averaging over the correct sequences for a particular code, not an average over a code ensemble as in the two preceding sections. For time invariant codes, the average is reduced to an average over all correct paths.

where $|\mathbf{x}-\mathbf{x}'| = \sqrt{\sum_{j=0}^{l-1}(x_j - x'_j)^2}$ is the Euclidean distance between the two signal sequences $\mathbf{x}$ and $\mathbf{x}'$.

For some nodes and sequences $\mathbf{x}$, the nearest neighbor is at distance $d_1 = \min_{\mathbf{x},\mathbf{x}'} |\mathbf{x} - \mathbf{x}'| = d_{free}$, for some it is at distance d_2, etc., up to some largest distance d_{n_d}. Further, let l_i be the minimum length of the error event that achieves d_i. If we collect all the node error probabilities and weight them according to their probability of occurrence p_i, we obtain from (17)

$$\begin{aligned} l_{\max} P_e &\geq \overline{P_2(\mathbf{x} \to \mathbf{x}')} \\ &= \sum_{i=1}^{n_d} p_i Q\left(d_i \sqrt{\frac{E_S}{2N_0}}\right), \qquad \sum_{i=1}^{n_d} p_i = 1, \end{aligned} \tag{18}$$

where p_i denotes the probability that the nearest incorrect sequence $\mathbf{x}'$ is at distance d_i, thus proving the theorem. Q.E.D.

Note that Theorem 2 is valid for time-invariant as well as for time-varying trellis codes, while we had to assume time-varying codes in the derivation of Theorem 1. We now combine these two theorems. Using the well-known approximation of the Q-function [7, page 83]

$$\frac{1}{\sqrt{2\pi}\gamma}(1 - \frac{1}{\gamma^2})e^{-\frac{\gamma^2}{2}} < Q(\gamma) < \frac{1}{\sqrt{2\pi}\gamma}e^{-\frac{\gamma^2}{2}}$$

in Theorem 2 and neglecting all terms $i > 1$, we obtain[5]

$$\begin{aligned} P_e &\geq \frac{p_1}{\sqrt{2\pi} l_1} \exp\left(-d_{free}^2 \frac{E_s/N_0}{4} - \ln\left(d_{free}\sqrt{\frac{E_s/N_0}{2}}\right) + \ln\left(1 - \frac{2}{d_{free}^2 E_s/N_0}\right)\right) \\ &= \frac{p_1}{\sqrt{2\pi} l_1} \exp\left(-d_{free}^2 \frac{E_s/N_0}{4}(1 + O(E_s/N_0))\right), \end{aligned} \tag{19}$$

where $O(E_s/N_0)$ is a quantity that goes to 0 as $E_s/N_0 \to \infty$.

Specializing Theorem 1 to AWGN-channels, we obtain

$$P_e \leq \left(\frac{1 - 2^{-R}}{2^{E(s)-R} - 1}\right)^{\frac{1}{s}} 2^{-\nu_m \frac{E(s)}{s}}, \quad E(s) > R, \tag{20}$$

where $0 < s \leq 1$ and

$$E(s) = -\log_2 \sum_{m=1}^{A} \sum_{p=1}^{A} p(a_m)p(a_p)e^{-s\frac{|a_m - a_p|^2}{4} E_s/N_0}.$$

We thus have an upper bound (20) and a lower bound (19) on the first event error probability of trellis codes on AWGN-channels, and therefore

$$\frac{p_1}{\sqrt{2\pi} l_1} \exp\left(-d_{free}^2 \frac{E_s/N_0}{4}(1 + O(E_s/N_0))\right) \leq \left(1 - 2^{-R}\right)^{\frac{1}{s}} \exp\left(-\frac{(\ln 2)\nu_m}{s} E(s) - \frac{1}{s}\ln\left(2^{E(s)-R} - 1\right)\right)$$

[5] This also allows us to set $l_{max} = l_1$.

where d_{free} is the normalized minimum free Euclidean distance of the *best* trellis code, since the upper bound is for the best trellis code (Theorem 1). We may take the natural logarithm of both sides to obtain

$$\begin{aligned}
-\ln(\sqrt{2\pi}l_1/p_1) - d_{free}^2\frac{E_s/N_0}{4}(1+O(E_s/N_0)) \;&\leq\; -\frac{(\ln 2)\nu_m}{s}E(s) - \frac{1}{s}\ln\left(2^{E(s)-R}-1\right) \\
&+\; \frac{1}{s}\ln\left(1-2^{-R}\right)
\end{aligned}$$

$$\begin{aligned}
d_{free}^2(1+O(E_s/N_0)) \;&\geq\; \frac{4(\ln 2)\nu_m}{sE_s/N_0}E(s) + \frac{4}{sE_s/N_0}\ln\left(2^{E(s)-R}-1\right) \\
&-\; \frac{4}{sE_s/N_0}\ln\left(1-2^{-R}\right) - \frac{4\ln(\sqrt{2\pi}l_1/p_1)}{E_s/N_0}.
\end{aligned}$$

For simplicity, let us denote $\frac{sE_s/N_0}{4}$ by α. Then we obtain

$$d_{free}^2(1+O(E_s/N_0)) \geq \frac{(\ln 2)\nu_m E(\alpha)}{\alpha} + \frac{\Theta\left(E(\alpha)\right)}{\alpha} - \frac{4\ln(\sqrt{2\pi}l_1/p_1)}{E_s/N_0}, \tag{21}$$

where

$$\begin{aligned}
\Theta\left(E(\alpha)\right) \;&=\; \ln\left(2^{E(\alpha)-R}-1\right) - \ln\left(1-2^{-R}\right) \\
E(\alpha) \;&=\; -\log_2\sum_{m=1}^{A}\sum_{p=1}^{A}p(a_m)p(a_p)e^{-\alpha|a_m-a_p|^2}, \\
0 \;&\leq\; \alpha \leq \frac{E_s/N_0}{4}.
\end{aligned}$$

We can now obtain a lower bound on the minimum free Euclidean distance d_{free} of the best code by letting $E_s/N_0 \to \infty$ in (21). This gives us the same bound derived in a different fashion by Rouanne and Costello [8], i.e.,

$$d_{free}^2 \geq \max_{\alpha\geq 0}\left(\frac{\ln(2)\nu_m E(\alpha)}{\alpha} + \frac{\Theta(E(\alpha))}{\alpha}\right), \quad E(\alpha) > R. \tag{22}$$

On the other hand, (20) can be written as

$$P_e \leq 2^{-\nu_m E_{ex}}, \tag{23}$$

where from the definition of α the *expurgated exponent* E_{ex} is given by

$$E_{ex} \triangleq \max_{0<\alpha\leq\frac{E_S/N_0}{4}}\left(\frac{E(\alpha)}{\alpha} + \frac{\Theta(E(\alpha))}{(\ln 2)\nu_m\alpha}\right)\frac{E_S/N_0}{4}, \quad E(\alpha) > R. \tag{24}$$

If the maximizing value of α in (22), $\alpha_{\max}$, is smaller then $(E_s/N_0)/4$, maximizing the minimum free Euclidean distance is the same as maximizing the expurgated error exponent E_{ex}.

For large ν_m, the contribution of the term $\Theta(E(\alpha))$ in (22) and (24) becomes negligible, and we may form the *asymptotic expurgated error exponent*

$$E_{ex}^{\infty} \triangleq \max_{0<\alpha\leq\frac{E_S/N_0}{4}} \left(\frac{E(\alpha)}{\alpha}\frac{E_S/N_0}{4}\right), \quad E(\alpha) > R, \tag{25}$$

and the bound of (22) becomes

$$\frac{d_{free}^2}{(\ln 2)\nu_m} = E_{ex}^{\infty}\frac{4}{E_S/N_0} = \frac{E(\alpha_{\max})}{\alpha_{\max}}, \tag{26}$$

where $\alpha_{\max}$ is the value of α which maximizes (25) and d_{free}^2/ν_m is the *asymptotic distance growth rate.* If $\alpha_{\max} \leq (E_S/N_0)/4$, then a signal constellation that maximizes the bound on the free distance will also maximize the expurgated error exponent. If, however, $\alpha_{\max} \geq (E_s/N_0)/4$, then the error bound (23) reduces to

$$P_e \leq \frac{1-2^{-R}}{2^{R_0-R}-1}2^{-\nu_m R_0}; \quad R_0 > R, \tag{27}$$

where

$$R_0 \triangleq -\log_2 \sum_{m=1}^{A}\sum_{p=1}^{A} p(a_m)p(a_p)e^{-\frac{|a_m-a_p|^2}{4}E_s/N_0} \tag{28}$$

is the cutoff-rate of the constellation on an AWGN-channel, and no expurgated error bound exists.

Maximizing E_{ex}^{∞} is the same as maximizing the function $E(\alpha)/\alpha$, which is accomplished easily with the help of the following lemma, which is proved in the appendix.

Lemma 3: *$E(\alpha)/\alpha$ is a monotonically decreasing function of α.*

Since $E(\alpha)/\alpha$ is a monotonically decreasing function of α, (25) achieves its supremum at the smallest value α such that $\alpha > 0$ and $E(\alpha) > R$. Since $E(\alpha)$, on the other hand, is a monotonically increasing function of α, $\alpha_{\max}$ is the smallest value of α such that $E(\alpha) > R$ and is given by the implicit equation

$$E(\alpha_{\max}) = R = -\log_2\left(\sum_{m=1}^{A}\sum_{p=1}^{A} p(a_m)p(a_p)e^{-\alpha_{\max}|a_m-a_p|^2}\right). \tag{29}$$

VI. A Geometric Construction

We now show how E_{ex}^{∞} and $\alpha_{\max}$ can be constructed from a graph of the cutoff rate R_0. As an example consider the 8-PSK constellation whose cutoff-rate R_0 in bits/signal is shown in Figure 5 (dotted line). When $E_s/N_0 > 4\alpha_{\max}$, (25) implies that E_{ex}^{∞} is a linear function of E_s/N_0 and, as can be seen from (29), its slope $E(\alpha_{\max})/4\alpha_{\max}$ depends only on the rate R for a fixed constellation. As $E_s/N_0 \to 4\alpha_{\max}$ from above, $E_{ex}^{\infty} \to R_0$. The higher the available energy, i.e., the larger E_s/N_0 is for a particular R, the larger E_{ex}^{∞} will be. In Figure

5, E_{ex}^{∞} is shown as solid lines over the range where (23) exists for several values of R. For a code with a larger value of R, the expurgated exponent grows more slowly with E_s/N_0 and a larger E_s/N_0 is required for the expurgated bound to exist.

With these preliminaries, E_{ex}^{∞} as well as the asymptotic distance growth rate, can easily be constructed from a graph of the cutoff-rate R_0. This construction is also illustrated in Figure 5.

Construction of the asymptotic expurgated error exponent E_{ex}^{∞} from the cutoff-rate R_0:

1. Choose the value of the code rate R. The cutoff-point is the intersection of a line a distance R above and parallel to the E_s/N_0-axis with the cutoff-rate curve. The x-value of the cutoff-point is $4\alpha_{\max}$.

2. Draw a straight line g through the origin of the graph and the cutoff-point.

3. The expurgated exponent for any $E_s/N_0 > 4\alpha_{\max}$ is the y-value of g at that value of E_s/N_0.

The asymptotic bound on d_{free}^2 from (26) is $4(\ln 2)\nu_m$ times the slope of g.

We should note the importance of R_0 at this point. If a constellation C_1 has a higher value of R_0 than constellation C_2 for some range of the signal-to-noise ratio E_S/N_0, then is evident from the above construction that trellis codes using constellation C_1, at a rate R such that $4\alpha_{\max}$ (the x-value of its cutoff-point) falls into that range, will have a larger expurgated error exponent E_{ex} as well as a larger asymptotic bound on the achievable free Euclidean distance d_{free} than trellis codes using constellation C_2. The merit of a constellation in conjunction with trellis codes can therefore be judged on the basis of its cutoff-rate R_0, and it is not necessary to evaluate either the expurgated bound or the bound on the minimum free Euclidean distance.

A constellation can now be optimized for Euclidean distance as well as event error probability by optimizing its cutoff-rate. Consider the upper envelope of the cutoff-rate curves for a set of possible signal constellations. Then using the above construction, the desired code rate R determines the constellation with the best cutoff-rate. This constellation then optimizes the Euclidean distance and the event error probability for this code rate R.

As an example of constellation optimization we have numerically optimized a pulse amplitude modulation (PAM) constellation with 8 signal points in Figure 6. It is interesting to see that for very small signal-to-noise ratios, $E_S/N_0 \leq 1$dB, the resulting constellation is in fact only 2-valued (BPSK). For larger E_S/N_0, successively more signal points move away from the clusters to form higher-sized constellations. At values of $E_S/N_0 > 13$dB, the constellation with uniform spacing (8-PAM) becomes optimal.

This optimization gives a cutoff-rate gain of up to a factor 2 (3dB) in E_S/N_0, as shown in Figure 7. This may be important for the construction of trellis codes for very low E_S/N_0 applications. It also confirms the well accepted observation that small-sized constellations are preferable for small values of E_S/N_0. The optimization of a PAM constellation with 4 signal points gives similar behavior, with much smaller gains in E_S/N_0, however.

It is not hard to show that the regular, unit-energy constrained, 2-dimensional constellation with 4 signal points (QPSK) is optimal in the above sense for all values of E_S/N_0. We have further observed numerically that the corresponding optimal circular constellation with 8 signal points is also regularly spaced (uniform 8-PSK).

VII. Conclusions

We have presented an expurgated bound on the first event error probability of trellis coded modulation on AWGN-channels. The asymptotic form of this bound is equivalent to known bounds on the minimum free Euclidean distance. The expurgated form of the bound gives, however, more information since it does not require an infinite signal-to-noise ratio to evaluate. The expurgated bound is a linear function of the signal-to-noise ratio and a simple construction, based on R_0, has been presented. The bound can also be used as a means of comparing different signal constellations.

VIII. Acknowledgement

The authors wish to thank Dr. Marc Rouanne for the many helpful discussions and his continued interest.

IX. Appendix: Proof of Lemma 3

We must show that $\frac{E(\alpha)}{\alpha}$ is a monotonically decreasing function of α. Let $F(\alpha) = \frac{E(\alpha)}{\alpha}$. Then

$$\begin{aligned}
F(\alpha+\varepsilon) &= \frac{1}{\alpha+\varepsilon}\left(-\log_2 \sum_{m=1}^{A}\sum_{p=1}^{A} p(\underline{a}_m)p(\underline{a}_p)\exp\left((\alpha+\varepsilon)\left|\underline{a}_m - \underline{a}_p\right|^2\right)\right) \\
&= \frac{1}{\alpha}\frac{\alpha}{\alpha+\varepsilon}\left(-\log_2 \sum_{m=1}^{A}\sum_{p=1}^{A} p(\underline{a}_m)p(\underline{a}_p)\exp\left((\alpha+\varepsilon)\left|\underline{a}_m - \underline{a}_p\right|^2\right)\right) \\
&= -\frac{1}{\alpha}\log_2 \left(\sum_{m=1}^{A}\sum_{p=1}^{A} p(\underline{a}_m)p(\underline{a}_p)\exp\left((\alpha+\varepsilon)\left|\underline{a}_m - \underline{a}_p\right|^2\right)\right)^{\frac{\alpha}{\alpha+\epsilon}}.
\end{aligned}$$

We now use Jensen's inequality (see, e.g., [12, appendix B]) for the special case $\overline{X^\beta} \le \overline{X}^\beta$, with $\beta < 1$, where the overbar denotes expectation, and obtain

$$\begin{aligned}
F(\alpha+\varepsilon) &\le -\frac{1}{\alpha}\log_2 \left(\sum_{m=1}^{A}\sum_{p=1}^{A} p(\underline{a}_m)p(\underline{a}_p)\exp\left(\frac{\alpha(\alpha+\varepsilon)}{(\alpha+\varepsilon)}\left|\underline{a}_m - \underline{a}_p\right|^2\right)\right) \\
&\le -\frac{1}{\alpha}\log_2 \sum_{m=1}^{A}\sum_{p=1}^{A} p(\underline{a}_m)p(\underline{a}_p)\exp\left(\alpha\left|\underline{a}_m - \underline{a}_p\right|^2\right) \\
&\le F(\alpha).
\end{aligned}$$

Thus $F(\alpha)$ is monotonically decreasing, which proves the lemma. Q.E.D.

References

[1] G. Ungerboeck, "Trellis-Coded Modulation with Redundant Signal Sets Part I: Introduction", *IEEE Communications Magazine*, Vol. 25, No.2, pp. 5–11, February 1987.

[2] G. Ungerboeck, "Trellis-Coded Modulation with Redundant Signal Sets Part II: State of the Art", *IEEE Communications Magazine*, Vol. 25, No.2, pp. 12–21, February 1987.

[3] G. Ungerboeck, "Channel Coding with Multilevel/Phase Signals", *IEEE Trans. Inform. Theory*, Vol. IT-28, No.1, pp. 55–67, January 1982.

[4] G. D. Forney, Jr., "Coset Codes–Part I: Introduction and Geometrical Classification", *IEEE Trans. Inform. Theory.,* Vol. IT-34, No.5, pp. 1123–1151, September 1988.

[5] G. D. Forney, Jr., "Coset Codes–Part II: Binary Lattices and Related Codes", *IEEE Trans. Inform. Theory.,* Vol. IT-34, No.5, pp. 1152–1187, September 1988.

[6] D. Divsalar and M. K. Simon, "Combined Trellis Coding with Asymmetric Modulations", *IEEE Trans. on Commun.*, Vol. COM-35, No.2, pp. 130–141, February 1987.

[7] J. M. Wozencraft and I. M. Jacobs, *Principles of Communication Engineering*, Wiley, New York, 1965.

[8] M. Rouanne and D. J. Costello, Jr., "A Lower Bound on the Minimum Euclidean Distance of Trellis-Coded Modulation Schemes", *IEEE Trans. Inform. Theory.,* Vol IT-34, No.5, pp. 1011–1020., September 1988.

[9] G. D. Forney, "Convolutional Codes I: Algebraic Structure", *IEEE Trans. Inform. Theory*, Vol. IT-16, No.6, pp. 720–738, November 1970.

[10] A. J. Viterbi and J. K. Omura, *Principles of Digital Communication and Coding*, McGraw-Hill, New York, 1979.

[11] R. G. Gallager, *Information Theory and Reliable Communication*, Wiley, New York, 1968.

[12] R. J. McEliece, *The Theory of Information and Coding*, Encyclopedia of Mathematics and its Applications, Volume 3, Addison-Wesley, London, 1977.

Figure Captions

Figure 1: Trellis coded modulation using quadrature modulation.

Figure 2: A correct and incorrect path pair through a trellis.

Figure 3: Cutoff-rates for different signal constellations for the AWGN-channel.

Figure 4: Some error paths diverging from the correct path at node j.

Figure 5: The cutoff-rate R_0 and the expurgated error exponent of 8-PSK.

Figure 6: Optimized 4-PAM and 8-PAM constellations for several rates R.

Figure 7: Optimized QPSK and 8-PSK constellations for several rates R.

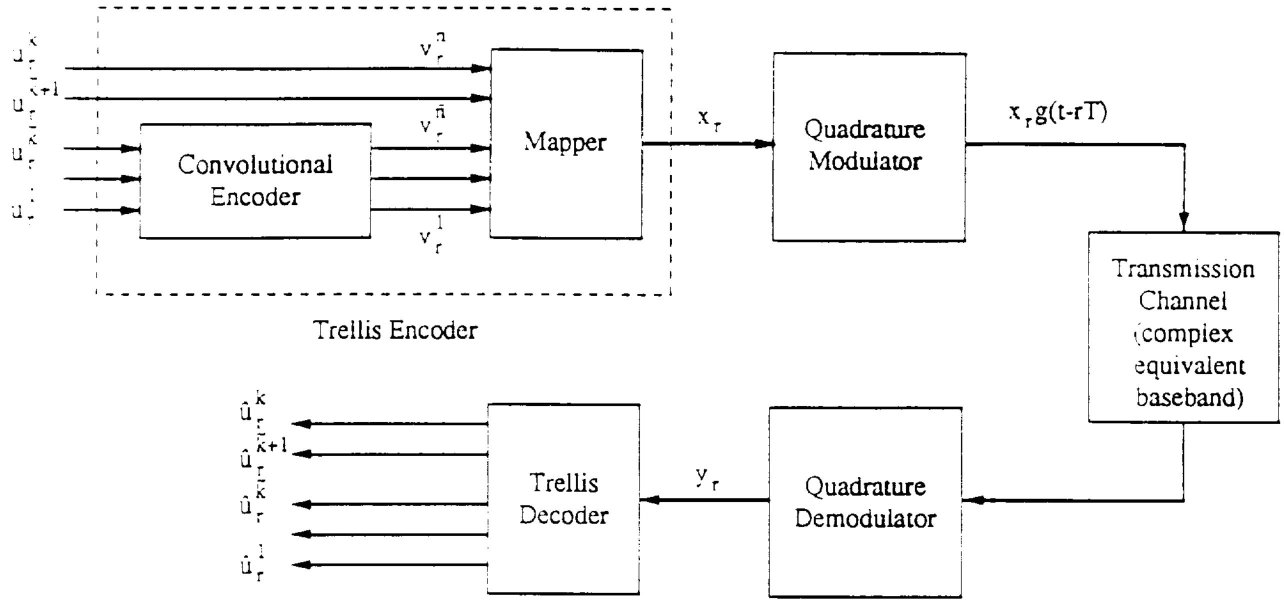


Figure 1: Trellis coded modulation using quadrature modulation.

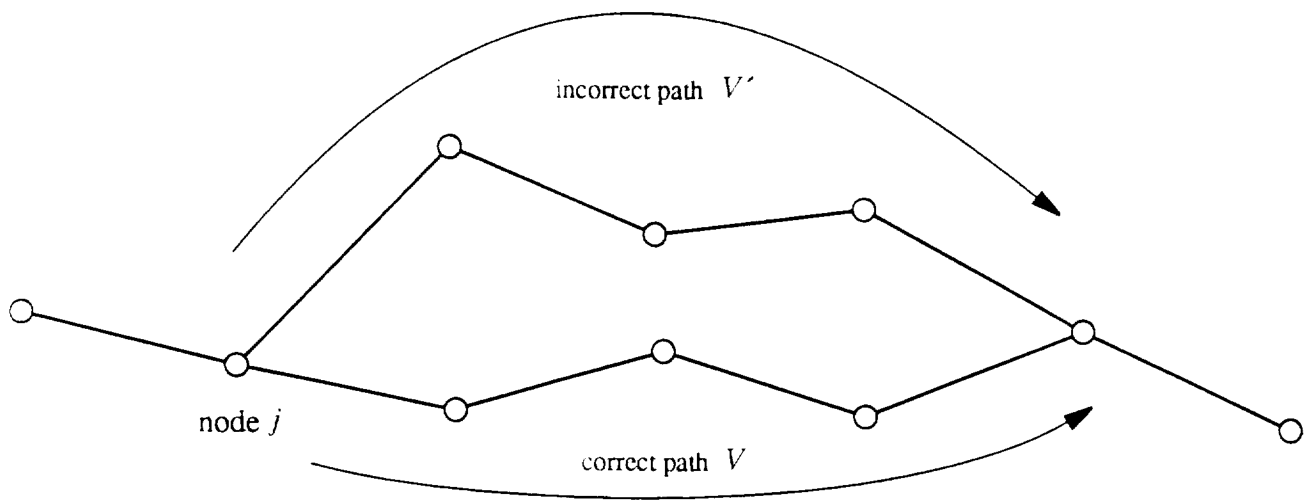


Figure 2: A correct and incorrect path pair through the trellis.

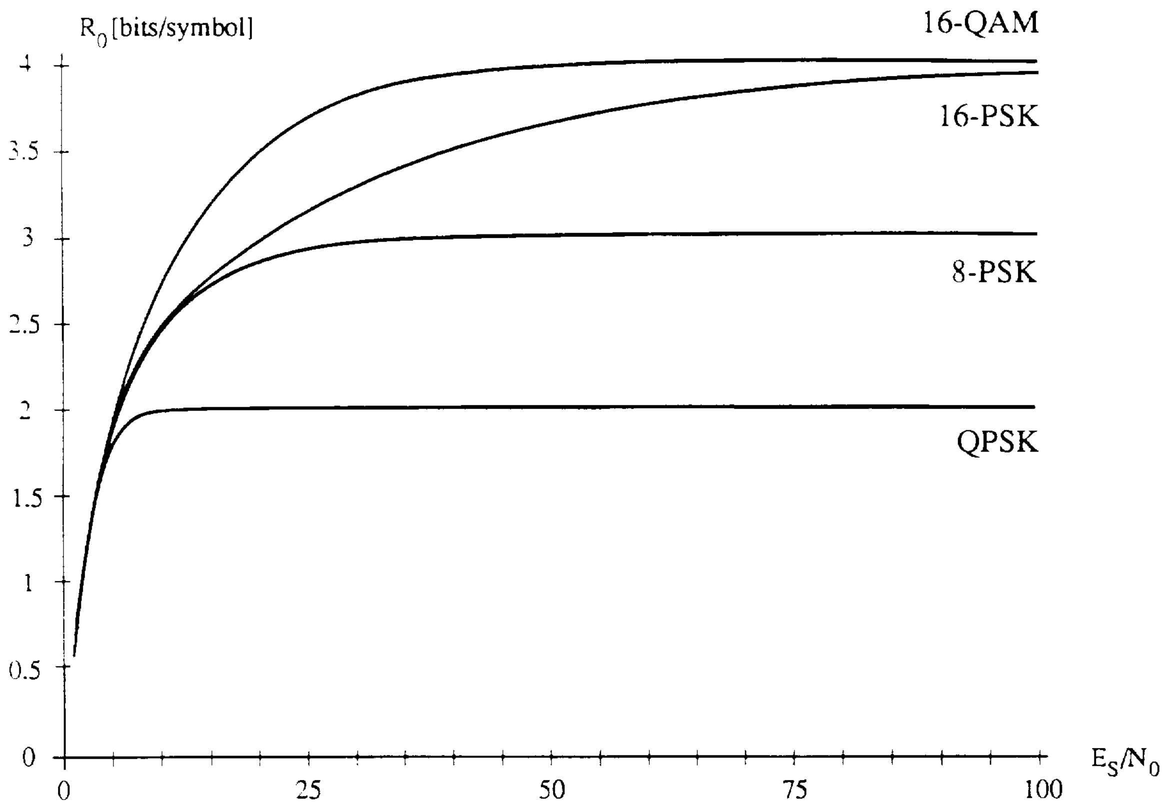


Figure 3: Cutoff-rates for different signal constellations for the AWGN-channel.

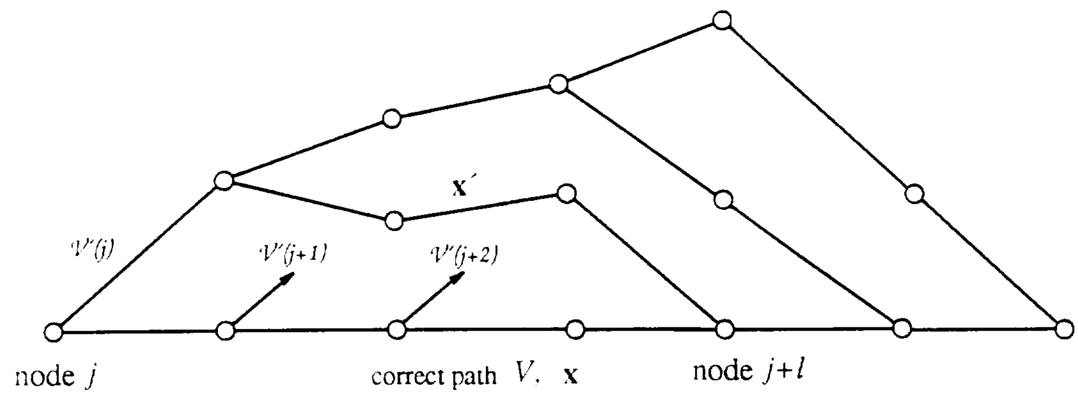


Figure 4: Some error paths diverging from the correct path at node j.

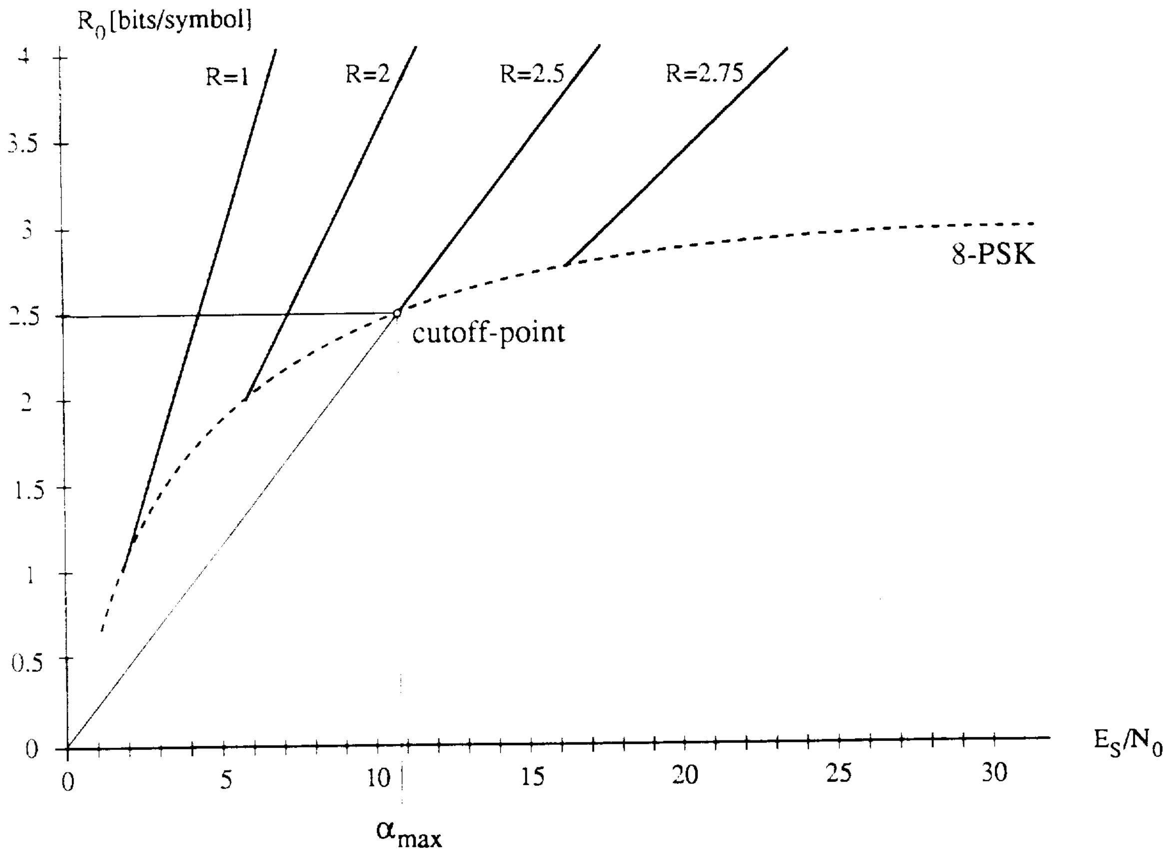


Figure 5: The cutoff-rate and the expurgated error exponent of 8-PSK.

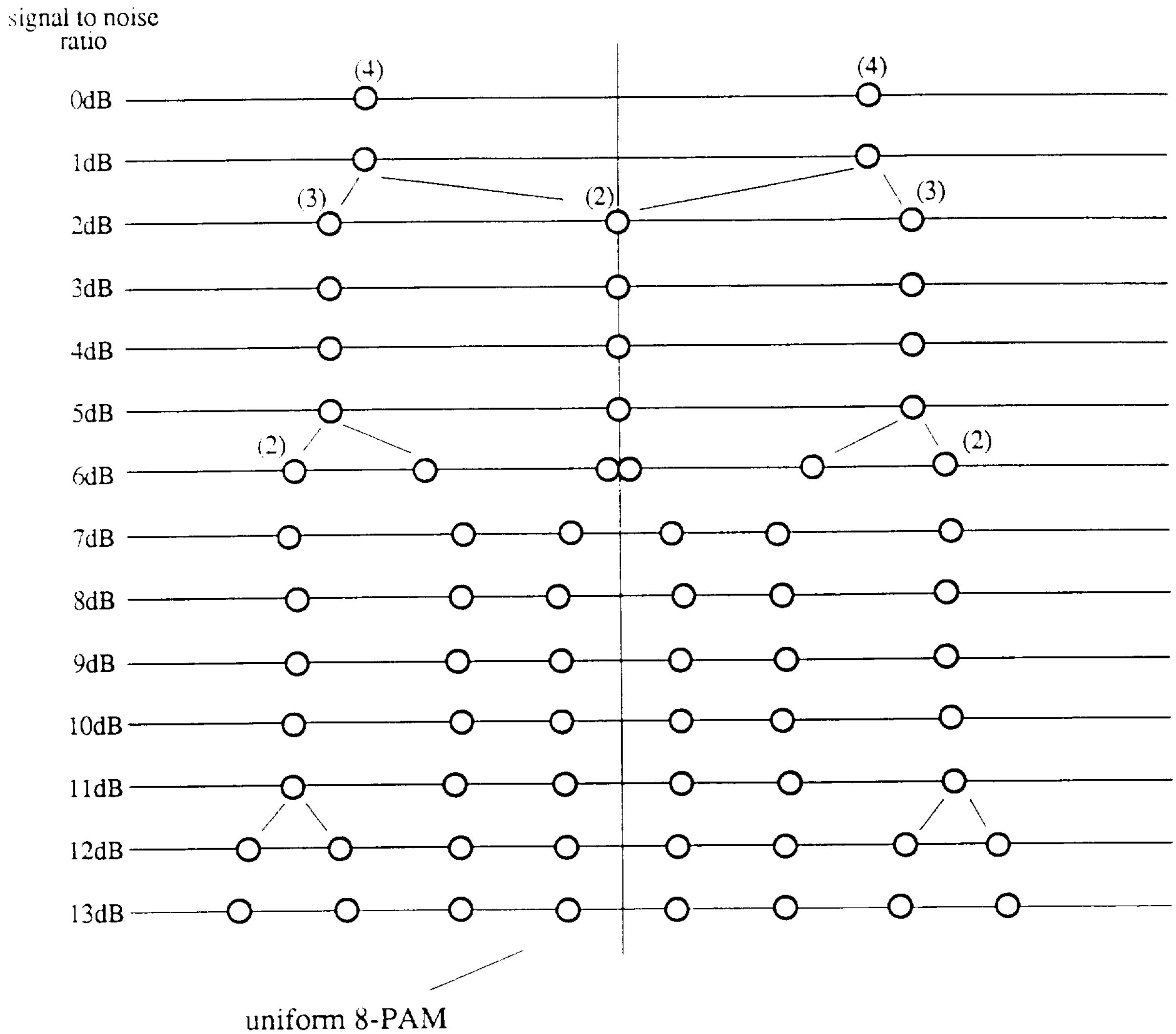


Figure 6: Optimized 8-PAM constellation. The numbers on the signal points are their multiplicity if different from 1.

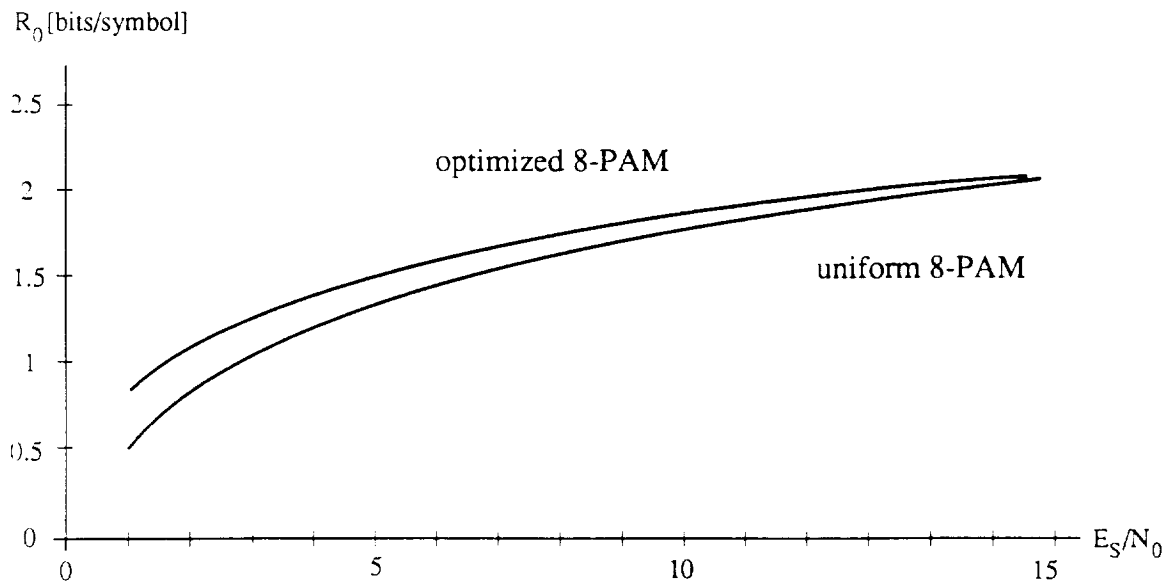


Figure 7: Cutoff-rates of the uniform 8-PAM and the optimized 8-PAM constellation.

Appendix B

An Upper Bound on the Symbol Error Rate for Convolutional and Trellis Codes

An Upper Bound on the Symbol Error Rate for Convolutional and Trellis Codes*

Lance C. Perez
and
Daniel J. Costello, Jr.

Department of Electrical and Computer Engineering
University of Notre Dame
Notre Dame, Indiana 46556

Presented at the 1990 International
Symposium on Information Theory

January 1990

*This work was supported in part by NASA Grants NAG5-557 and NGT-70109 and NSF Grant NCR89-03429.

Introduction

- In concatenated coding systems using Reed-Solomon (RS) outer codes over $GF(2^b)$ and ideal symbol interleaving between the inner and outer code, the system bit error rate (BER) is closely approximated by

$$P_b = \frac{2t+1}{N} \sum_{i=t+1}^{N} \binom{N}{i} P_s^i (1-P_s)^{N-i} \tag{1}$$

t is the error correcting capability of the RS code

where P_s is the symbol error rate (SER) out of the inner decoder.

- Thus, for a given outer code the performance measure of the inner code is the SER and *not* the BER.

- For convolutional and trellis inner codes, simulation is generally used to obtain P_s. The byte size of the outer code symbol requires a very large number of bits to be simulated for statistically valid points.

- Goal: Find an analytic method for determining P_s for convolutional and trellis codes. *(Onyszchuk and McEliece for $1/n$ convolutional codes)*

Conceptual Motivation

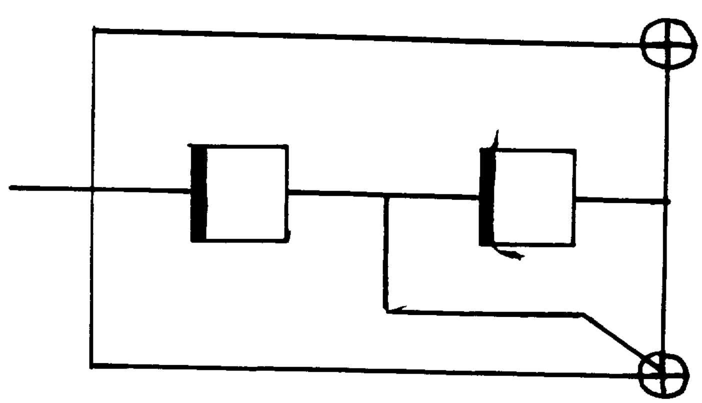
- Reinterpret the union bound on the probability of bit error for a convolutional code given by

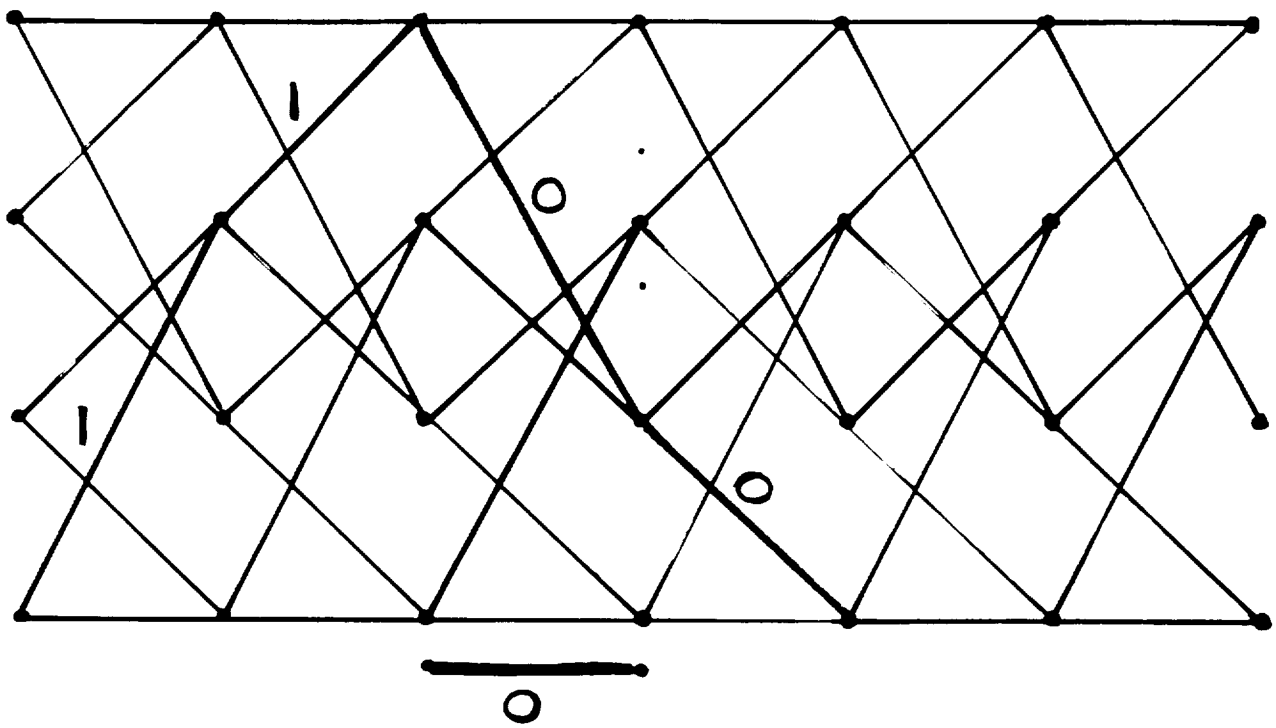
$$P_b = \frac{1}{k} \sum_{d=d_{free}}^{\infty} B_d P_d \qquad (2)$$

where B_d is the total number of nonzero information bits on all weight d paths and P_d is the two codeword error probability.

- Traditionally, this bound is developed by considering a truncated trellis as a block code and then computing the average number of information bit errors per decoded information block. Another derivation of this bound is useful.

- Example: $r = 1/2$, $m = 2$, convolutional code with the following feedforward encoder realization and trellis.





- An error event of weight d occurs with probability P_d. By simple counting, this error event can cause the information bit on the j^{th} branch to be in error in precisely $B_d = 2$ ways. Where B_d is the number of nonzero informaton bits on the incorrect path.
- Thus, the BER due to this error event, denoted P_{b_d}, is

$$P_{b_d} = B_d P_d \tag{3}$$

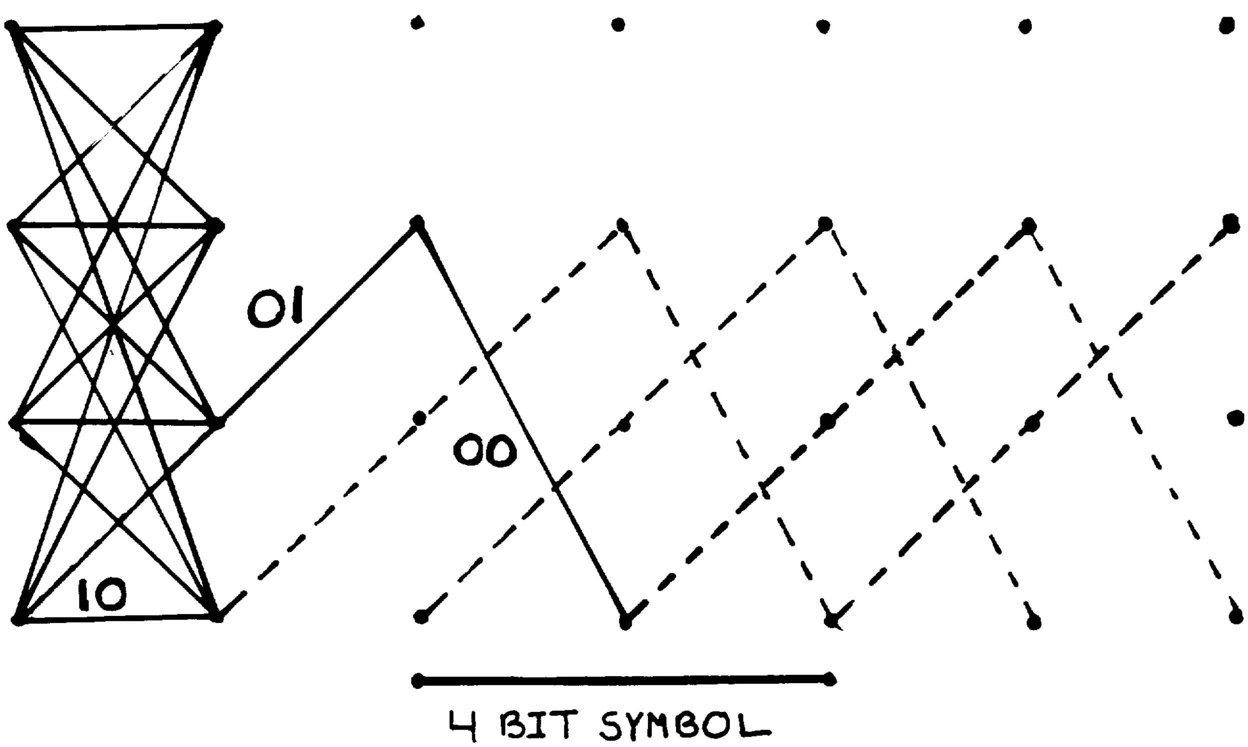
- Summing over all possible error events yields

$$P_b = \sum_{d=d_{free}}^{\infty} B_d P_d \tag{4}$$

- The counting technique used to determine the upper bound on P_b can be extended to bound the SER out of the inner decoder.

Bound on the Symbol Error Rate

- Example: $r = 2/3$, $m = 1$ convolutional code with a feedforward encoder and the following trellis and with 4-bit symbols for the outer RS code.

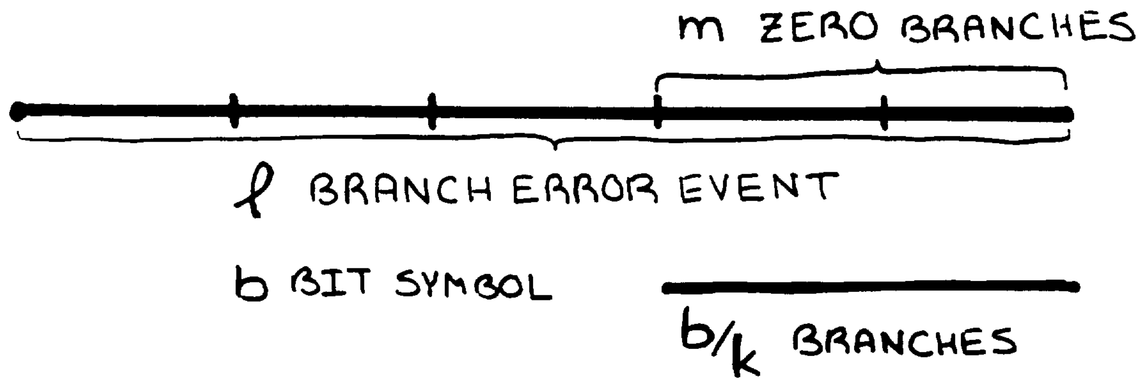


Assuming that symbol boundaries are always aligned with trellis nodes, the error event shown can cause the particular 4-bit symbol to be in error in 3 ways, each occuring with probability P_d.

- Thus, the probability of symbol error due to this error event is

$$P_{s_{d,l}} = 3P_d \tag{5}$$

- An error event of length l branches and weight d, can cause an error in <u>at most</u> $(b/k + l - m - 1)$ ways, each occuring with probability P_d.



- The $-m$ term is due to the fact that in <u>feedforward</u> realizations, all error events end with m consecutive 0 branches.

- Summing over all error events of all lengths gives,

$$P_s \leq \sum_{d=d_{free}}^{\infty} \sum_{l=m+1}^{\infty} (b/k + l - m - 1) A_{d,l} P_d \tag{6}$$

where $A_{d,l}$ is the number weight d paths with length l. This can be simplified to

$$P_s \leq (b/k - m - 1) \sum_{d=d_{free}}^{\infty} A_d P_d + \sum_{d=d_{free}}^{\infty} L_d P_d \tag{7}$$

where

$$L_d = \sum_{l=m+1}^{\infty} l A_{d,l} \tag{8}$$

is the <u>total length in branches</u> of all weight d paths.

- Performance Factors:

1. Path multiplicity, A_d, is dominant.
2. Degree of byte orientation, b/k. *(Lin-nan Lee)*
3. Length of the error events. *(Simon and Divsalar)*

- In terms of the code transfer function

$$P_s \leq K(d_f)\left((b/k - m - 1)T(D,L,I) + \frac{\partial T(D,L,I)}{\partial L}\right)\Big|_{\substack{L=I=1 \\ D=\exp(-E_s/N_0)}} \tag{9}$$

- In systematic feedback encoder realizations, error events cannot end in all zeroes branches. Thus, the bound becomes

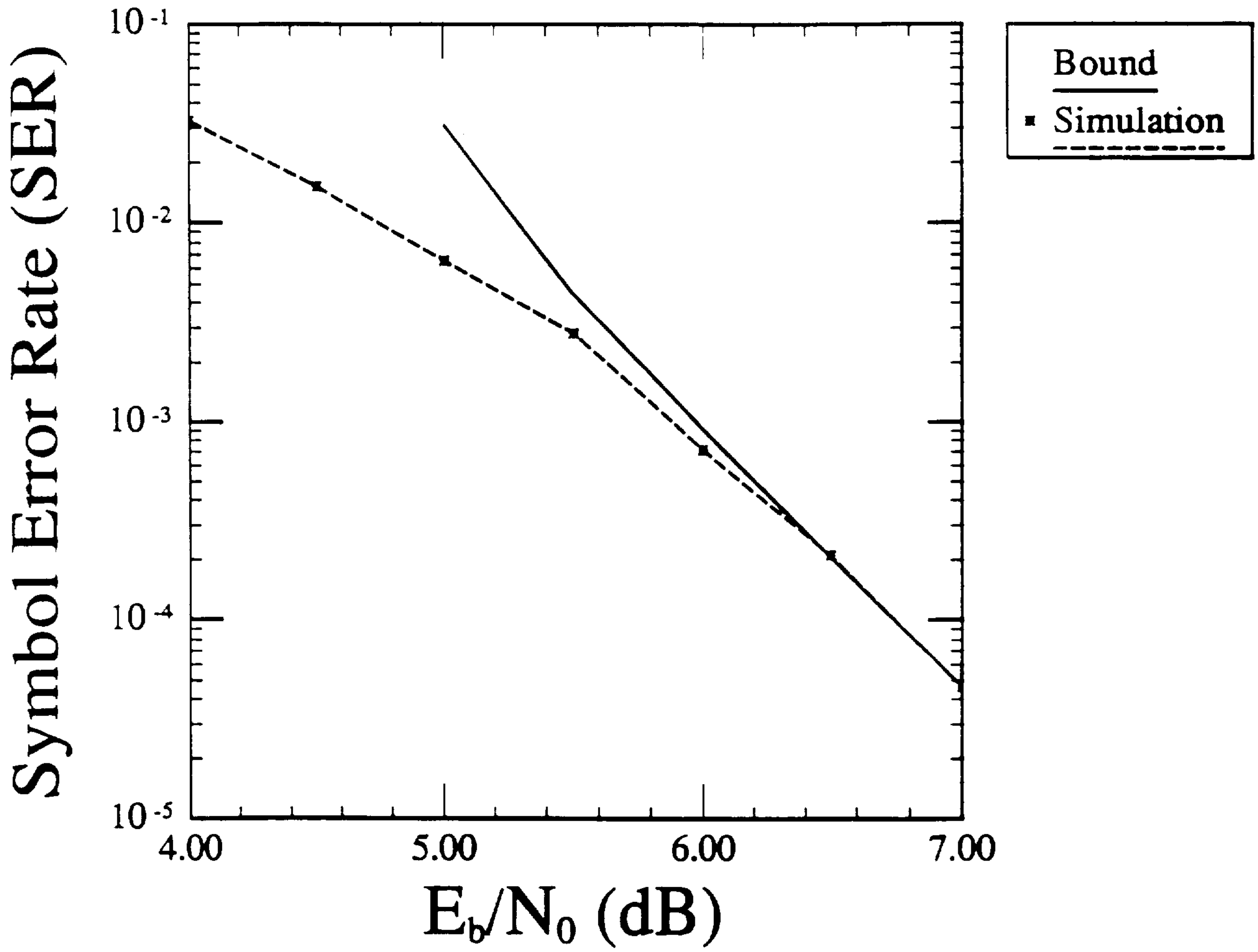
$$P_s \leq K(d_f)\left((b/k - 1)T(D,L,I) + \frac{\partial T(D,L,I)}{\partial L}\right)\Big|_{\substack{L=I=1 \\ D=\exp(-E_s/N_0)}} \tag{10}$$

Trellis Codes

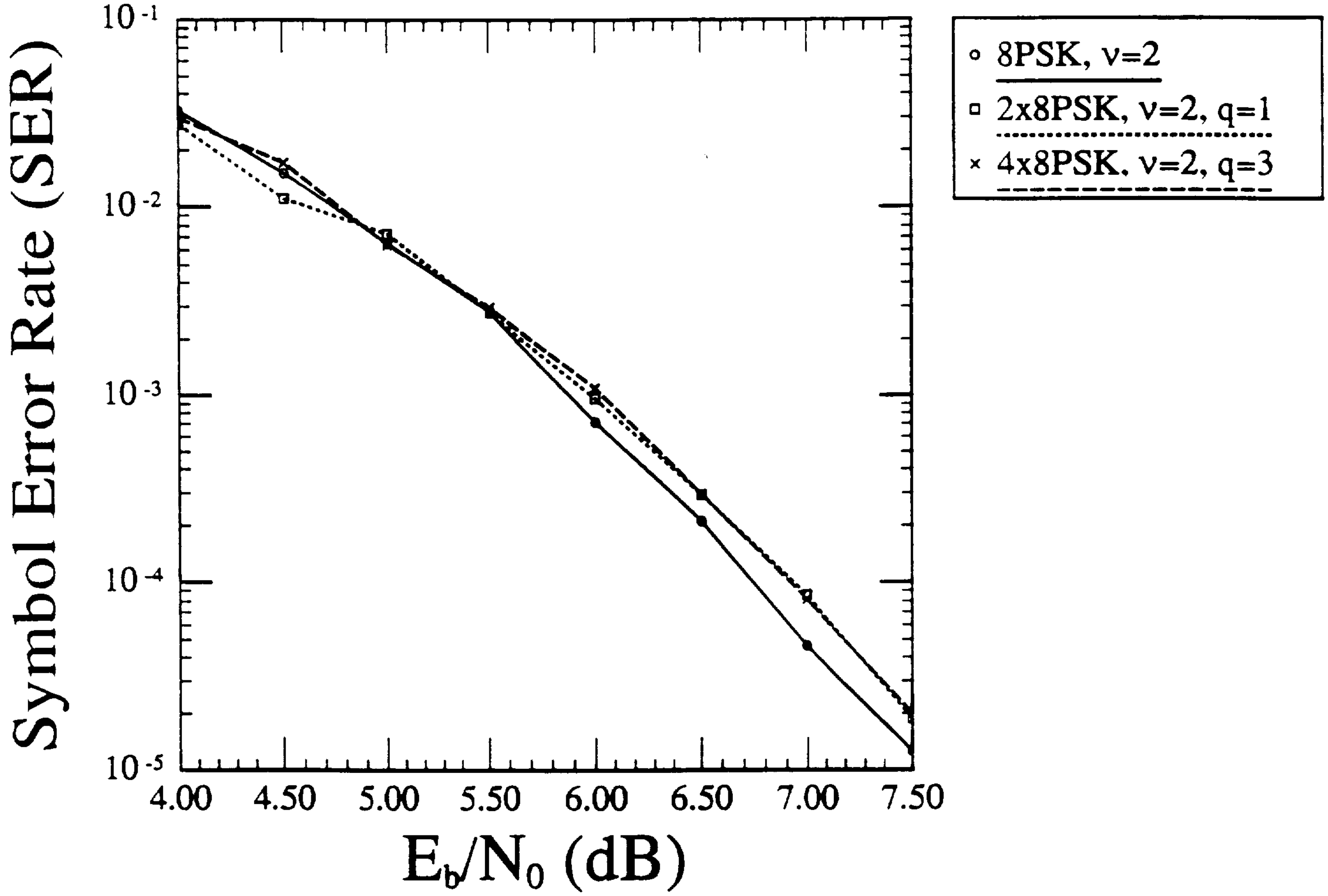
- For appropriate trellis codes, a bound on the SER can be obtained using the Zehavi and Wolf transfer function.
- For the LxMPSK codes constructed by Pietrobon, et.al. and Ungerboeck, systematic feedback encoders are used and the bound becomes

$$P_s \leq K(d_f)\left((b/k - 1)T(W,L,I) + \frac{\partial T(W,L,I)}{\partial L}\right)\Big|_{\substack{L=I=1 \\ W=\exp(-E_s/4N_0)}} \tag{11}$$

Simulation vs. Bound
Ungerboeck $\nu = 2$, 8PSK Code



Simulation Results
Multi-Dimensional 8PSK Trellis Codes



Conclusions

- An upper bound on the SER for convolutional/trellis codes can be obtained using a transfer function approach.
- Feedforward realizations of a particular code may perform better than the feedback realization of the same code in concatenated systems.
- For concatenated systems, it may be better to design the inner code to have a short d_{free} path, i.e. to design the d_{free} path to be a parallel transition.
- Multi-D trellis codes with byte oriented branches do not improve in SER compared to 2D Ungerboeck codes because of high path multiplicities and dense spectra.

Appendix C

A General Parity Check Equation for Rotationally Invariant Trellis Codes

A GENERAL PARITY CHECK EQUATION FOR ROTATIONALLY INVARIANT TRELLIS CODES[1]

Steven S. Pietrobon
Daniel J. Costello, Jr.

University of Notre Dame
Dept. of Elec. & Comp. Engr.
Notre Dame, IN 46556
U.S.A.

Gottfried Ungerboeck

IBM Zurich Research Laboratory
Säumerstrasse 4
CH-8803, Rüschlikon
Switzerland

Presented at the 1989 Workshop on Information Theory
Ithaca, New York,
U.S.A.
June 26, 1989

[1]This work was supported by NASA Grant NAG5-557.

Trellis Codes With Linear Parity Check Equations.

- For rate $k/(k+1)$ trellis codes, the *parity check equation* defines the relationship between the $k+1$ binary output sequences $y^0(D), y^1(D), \ldots, y^k(D)$.

- A *linear* parity check equation:

$$H^k(D)y^k(D) \oplus \cdots \oplus H^1(D)y^1(D) \oplus H^0(D)y^0(D) = 0(D)$$

where $H^i(D)$ = parity check polynomial of $y^i(D)$.
$0(D)$ = all zeros sequence.

- The *constraint length* (ν) of an encoder is the maximum degree of all $H^i(D)$, i.e.,

$$\nu = \max_{all\ i} \deg\ H^i(D)$$

- The *memory* (m) of an encoder is the number of delay elements required to implement an encoder.

- For linear codes it can be shown that $m = \nu$.
(See Forney, "Convolutional Codes I", *IEEE Trans. on Inform. Theory*, November 1970).

- The *integer representation* of $y^i(D)$, for $0 \leq i \leq k$, is defined as

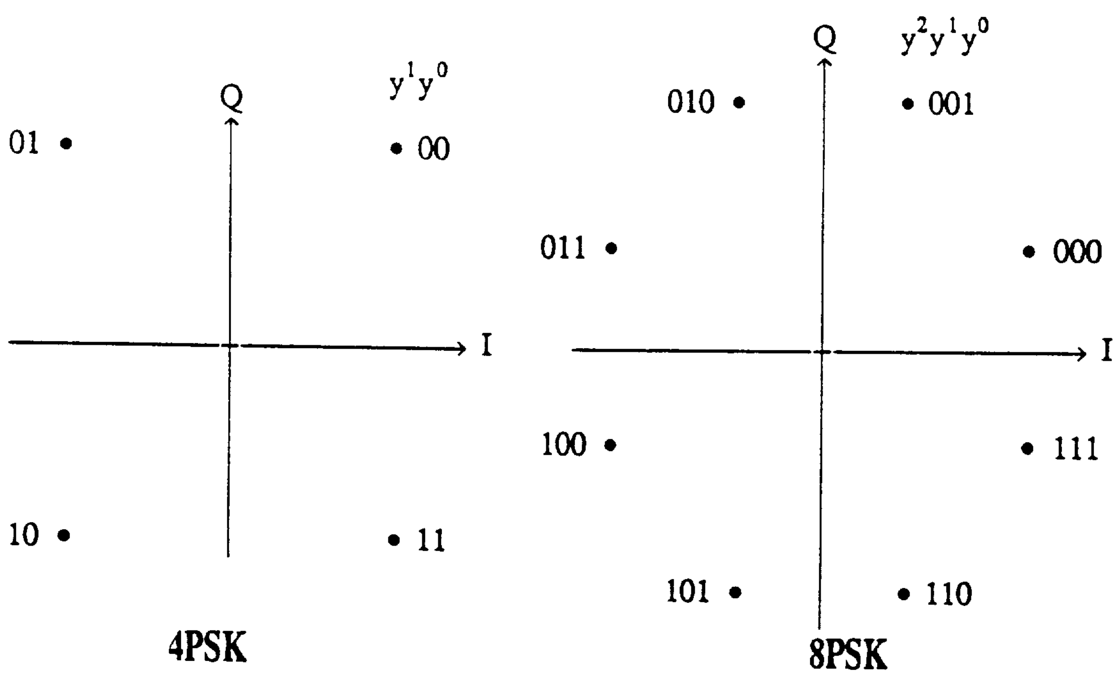
$$\begin{aligned} y(D) &= y^0(D) + 2y^1(D) + \cdots + 2^k y^k(D) \\ &= \sum_{i=0}^{k} 2^i y^i(D) \end{aligned}$$

- We define a naturally mapped signal set as a signal set such that a discrete phase rotation of the signal set produces a *rotated sequence* $y_r(D)$

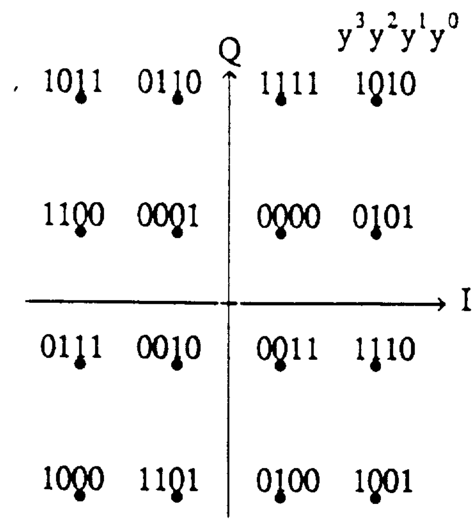
$$y_r(D) = y(D) + 1(D) \qquad (\text{mod M})$$

where $\quad 1(D) =$ all ones sequence
$\quad M = 2^{k+1}$

<u>Example:</u> MPSK



Special Case: Naturally Mapped 16 QAM



Here $\qquad y(D) = y^0(D) + 2y^1(D)$

with $\qquad y_r(D) = y(D) + 1(D) \quad$ **(mod 4),**

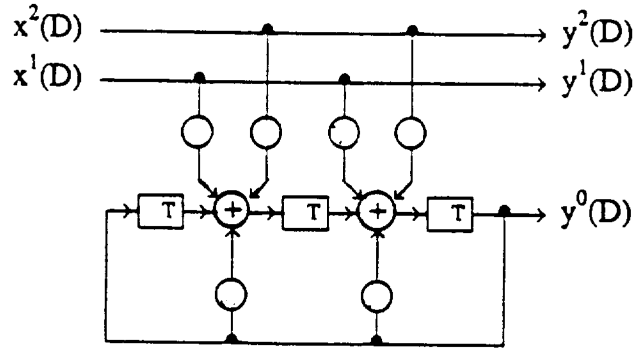
$y_r^2(D) = y^2(D)$ **and** $y_r^3(D) = y^3(D)$.

Systematic Encoding

For a systematic encoder we let

$$\begin{aligned} y^1(D) &= x^1(D) \\ y^2(D) &= x^2(D) \\ &\vdots \\ y^k(D) &= x^k(D) \end{aligned}$$

Example of Systematic Encoder with $\nu = 3$ and $k = 2$ (rate 2/3).



Effect of Phase Rotation on Linear Parity Check Equations

- With natural mapping we have

$$
\begin{aligned}
y_r^0 &= y^0 \oplus 1 = \overline{y^0} \\
y_r^1 &= y^1 \oplus y^0 \\
y_r^2 &= y^2 \oplus y^0 \cdot y^1 \\
&\vdots \\
y_r^k &= y^k \oplus \prod_{i=0}^{k-1} y^i
\end{aligned}
$$

- On a phase rotation the parity check equation becomes

$$
H^k(D)y_r^k(D) \oplus H^{k-1}(D)y_r^{k-1}(D) \oplus \cdots \oplus H^1(D)y_r^1(D) \oplus H^0(D)y_r^0(D) = 0(D)
$$

$$
\begin{aligned}
H^0(D)y_r^0(D) &= H^0(D)(y^0(D) \oplus 1(D)) \\
&= H^0(D)y^0(D) \oplus H^0(D)1(D) \\
&= H^0(D)y^0(D) \oplus E[H^0(D)](D)
\end{aligned}
$$

where $E[H^0(D)]$ is the modulo-2 number of non-zero terms in $H^0(D)$, e.g., $E[D^5 \oplus D^4 \oplus D^3 \oplus D^2] = 0$.

- **If** $E[H^0(D)] = 0$, **then** $H^0(D)y_r^0(D) = H^0(D)y^0(D)$.

$$\begin{aligned} H^1(D)y_r^1(D) &= H^1(D)(y^1(D) \oplus y^0(D)) \\ &= H^1(D)y^1(D) \oplus H^1(D)y^0(D) \\ &\neq H^1(D)y^1(D) \end{aligned}$$

- **Thus linear parity check equations are <u>not</u> phase transparent.**

A Parity Check Equation Not Affected by a Phase Rotation

- Assume that $E[H^0(D)] = 0$. Let

$$\boxed{z(D) = (D^a + (M-1)D^b)y(D) \qquad \textbf{(mod M)}}$$

where $\nu > a > b > 0$.

- On a phase rotation

$$\begin{aligned} z_r(D) &= (D^a + (M-1)D^b)y_r(D) && \textbf{(mod M)} \\ &= (D^a + (M-1)D^b)(y(D) + 1(D)) && \textbf{(mod M)} \\ &= (D^a + (M-1)D^b)y(D) + (D^a + (M-1)D^b)1(D) && \textbf{(mod M)} \end{aligned}$$

- Note that $D^i 1(D) = 1(D)$ for all integers i. Thus

$$\begin{aligned} z_r(D) &= z(D) + 1(D) + (M-1)(D) && \textbf{(mod M)} \\ &= z(D) + M(D) && \textbf{(mod M)} \\ &= z(D) \end{aligned}$$

- Thus all the bits in $z(D)$ are unaffected by a phase rotation.

- Note that the most significant bit of $z(D)$ is a function of all $y^i(D)$, satisfying the requirement that these bits are checked by the encoder.

- We have that

$$z(D) = z^0(D) + 2z^1(D) + \cdots + 2^k z^k(D)$$

and

$$H^0(D) = D^\nu \oplus h_0^{\nu-1} D^{\nu-1} \oplus \cdots \oplus h_0^2 D^2 \oplus h_0^1 D \oplus 1.$$

- We form the parity check equation

$$\boxed{z^k(D) \oplus h_{k-1}^z z^{k-1}(D) \oplus \cdots \oplus h_1^z z^1(D) \oplus H^0(D) y^0(D) = 0(D)}$$

$z^k(D)$ is always selected, since it checks all input bits (thus avoiding parallel transitions).
h_i^z are used to select other bits of $z(D)$.
$z^0(D)$ is not selected since it is a linear function of $y^0(D)$ (which is taken care of by $H^0(D)y^0(D)$ in the parity check equation).

- In implementing an encoder, need to determine $z^i(D)$ in terms of $y^0(D), y^1(D), \ldots, y^k(D)$.

- If $E[H^0(D)] = 1$, we let

$$\boxed{z(D) = (D^a + (M/2-1)D^b)y(D) \qquad \textbf{(mod M)}}$$

- With this form of $z(D)$ we have

$$\begin{aligned} z_r(D) &= (D^a + (M/2-1)D^b)(y(D) + 1(D)) && \textbf{(mod M)} \\ &= (D^a + (M/2-1)D^b)y(D) + 1(D) + (M/2-1)(D) && \textbf{(mod M)} \\ &= z(D) + (M/2)(D) && \textbf{(mod M)} \end{aligned}$$

- Thus we have $z_r^i(D) = z^i(D)$ for $0 \le i \le k-1$ (i.e., the first k least significant bits of $Z(D)$ are unaffected by a phase rotation) but

$$z_r^k(D) = z^k(D) \oplus 1(D).$$

- Since $z^k(D)$ is always selected, the $1(D)$ term generated by $z^k(D)$ will cancel the $1(D)$ term generated by $H^0(D)y^0(D)$.

- We can also have other forms of $z(D)$, as long as $z^k(D)$ checks all the bits in $y(D)$. For example (with $E[H^0(D)] = 0$),

$$z(D) = (D^a + 3D^b + 4D^c)y(D) \qquad \textbf{(mod 8)}.$$

Example: Rate 1/2 QPSK (M = 4)

- The input sequence

$$x(D) = x^1(D)$$

and the output sequence

$$y(D) = y^0(D) + 2y^1(D).$$

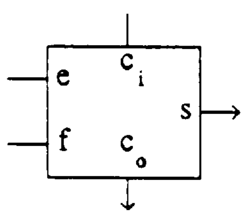
- We have

$$z(D) = (D^a + 3D^b)y(D) \qquad \textbf{(mod 4).}$$

- We need to express $z(D)$ in terms of $y^0(D)$ and $y^1(D)$:

$$\begin{aligned} z(D) &= (D^a + 3D^b)y(D) & \textbf{(mod 4)} \\ z(D) &= (D^a + D^b + 2D^b)(y^0(D) + 2y^1(D)) & \textbf{(mod 4)} \\ z^0(D) + 2z^1(D) &= (D^a + D^b)y^0(D) + 2((D^a + D^b)y^1(D) + D^b y^0(D)) \end{aligned}$$

- For a two bit binary adder



$$s = e \oplus f \oplus c_i$$

$$c_0 = e \cdot f \oplus c_i \cdot (e \oplus f)$$

- Thus

$$\begin{aligned} z^0(D) &= (D^a \oplus D^b)y^0(D) \qquad \textbf{(not used)} \\ z^1(D) &= (D^a \oplus D^b)y^1(D) \oplus D^b y^0(D) \oplus D^a y^0(D) \cdot D^b y^0(D) \\ &= (D^a \oplus D^b)y^1(D) \oplus \overline{D^a y^0(D)} D^b y^0(D) \end{aligned}$$

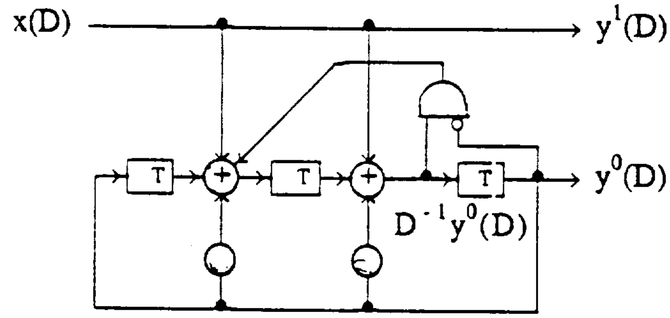
- The parity check equation becomes

$$(D^a \oplus D^b)y^1(D) \oplus \overline{D^a y^0(D)} \cdot D^b y^0(D) \oplus H^0(D)y^0(D) = 0(D)$$

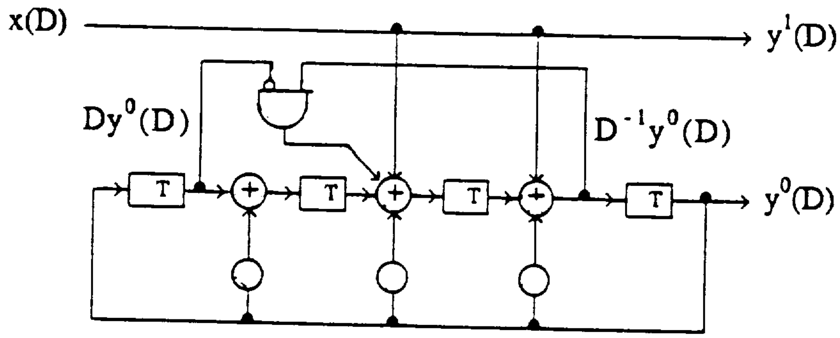
Example of Rate 1/2 Systematic Encoder With Feedback

- We have $\nu = 3, a = 2$, and $b = 1$, which gives the parity check equation

$$(D^2 \oplus D)y^1(D) \oplus \overline{D^2y^0(D)} \cdot Dy^0(D) \oplus (D^3 \oplus h_0^2 D^2 \oplus h_0^1 D \oplus 1)y^0(D) = 0(D)$$



- Example of rate 1/2 encoder with $\nu = 4, a = 3$, and $b = 1$.



- For $|a - b| > 2$ the encoder may not be minimal or $H^0(D)$ may need to be restricted.

Example: Rate 2/3 8PSK (M = 8)

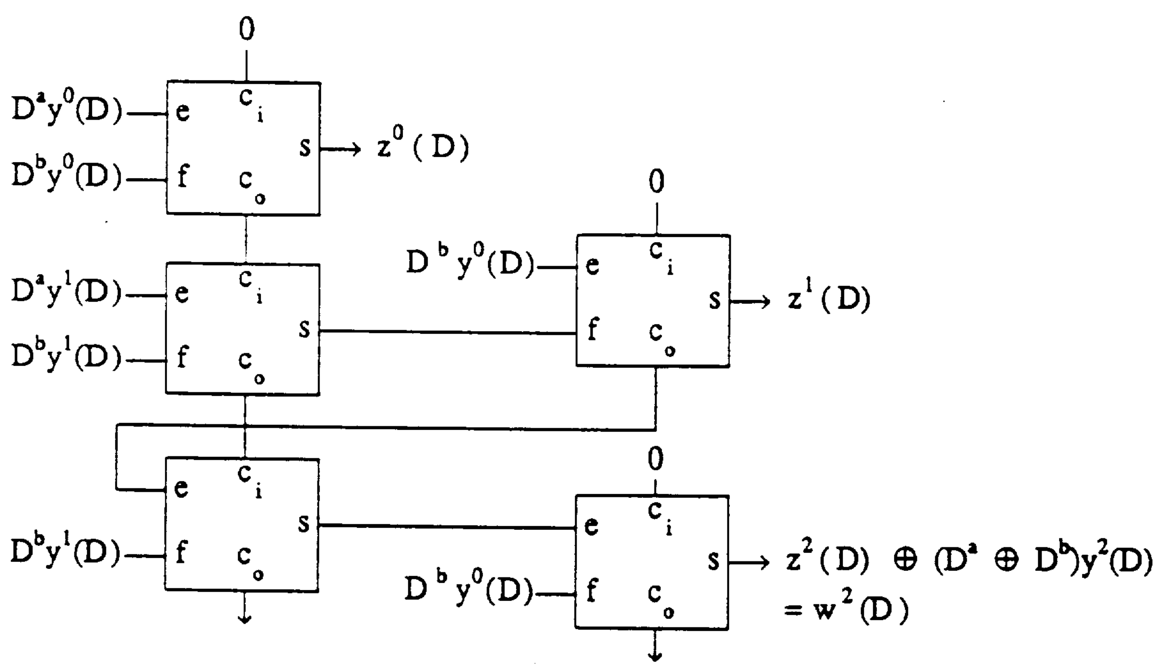
- We have

$$z(D) = (D^a + 7D^b)y(D) \qquad \textbf{(mod 8)}$$

- Expressing $z(D)$ in terms of $y^0(D), y^1(D)$, and $y^2(D)$

$$\begin{aligned} z(D) &= (D^a + D^b + 2D^b + 4D^b)(y^0(D) + 2y^1(D) + 4y^2(D)) \qquad \textbf{(mod 8)} \\ z(D) &= (D^a + D^b)y^0(D) + \\ &\quad 2((D^a + D^b)y^1(D) + D^b y^0(D)) \\ &\quad 4((D^a + D^b)y^2(D) + D^b y^1(D) + D^b y^0(D)) \qquad \textbf{(mod 8)} \end{aligned}$$

- Using two bit logic adders

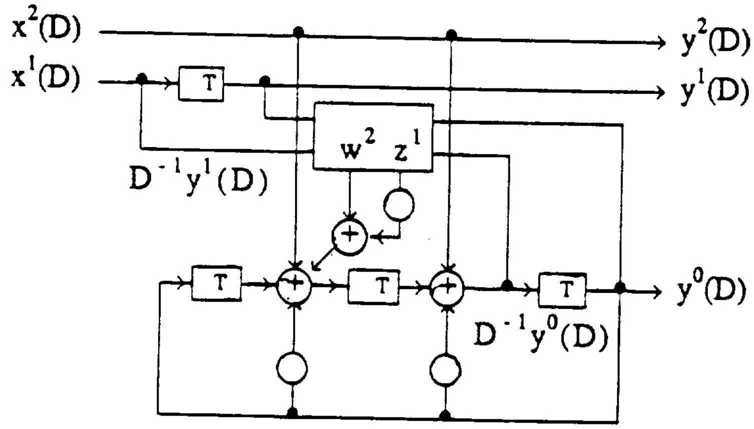


or

$$\begin{aligned} z^2(D) &= (D^a \oplus D^b)y^2(D) \oplus D^b(y^1(D) \oplus y^0(D)) \\ &\quad \oplus D^a y^1(D) \cdot D^b y^1(D) \oplus D^a y^0(D) \cdot D^b y^0(D) \cdot (D^a \oplus D^b)y^1(D) \\ &\quad \oplus D^b y^0(D) \cdot ((D^a \oplus D^b)y^1(D) \oplus D^a y^0(D) \cdot D^b y^0(D)) \end{aligned}$$

- Example of rate 2/3 encoder with $\nu = 3, a = 2$, and $b = 1$.
- Parity check equation:

$$(D^2 \oplus D)y^2(D) \oplus w^2(D) \oplus h_z^1 z^1(D) \oplus (D^3 \oplus h_0^2 D^2 \oplus h_0^1 D \oplus 1)y^0(D) = 0(D)$$



- Note that encoder is not minimal.

Conclusions

- Trellis codes based on linear parity check equations are not rotationally invariant.
- A general parity check equation for rotationally invariant trellis codes has been presented.
- A method of finding an encoder implementation for these codes has been given.
- Not all rotationally invariant codes are minimal. Rate $k/(k+1)$ codes with two or more checked bits are not minimal.
- Method can be applied to all signal sets with phase symmetry by appropriately mapping points in the signal set.
- Since codes are non-linear, a systematic code search involves searching all paths to find the free distance.

Appendix D

Trellis Coding Using Multi-Dimensional QAM Signal Sets

TRELLIS CODING USING MULTI-DIMENSIONAL QAM SIGNAL SETS*

by

Steven S. Pietrobon and Daniel J. Costello, Jr.

June 1989

Submitted to the

1990 IEEE International Symposium on Information Theory

Abstract

A method of finding good trellis codes with multi-dimensional (multi-D) QAM modulation is presented. Using the 16QAM signal set, 4-D, 6-D, and 8-D QAM signal sets are constructed which have good partition and phase rotational properties.

The good partition properties are achieved by the use of block codes and their cosets restricting each level in the multi-D mapping. The rotational properties are achieved through the use of a "naturally mapped" 16 QAM signal set. This signal set has the property that, of the four bits used to map the signal set, only two bits are affected by a 90° phase rotation. With an appropriate addition of the coset generators, the multi-D signal sets also have two mapping bits affected by a 90° phase rotation (the remaining bits being unaffected).

This implies that many good rate $k/(k+1)$ trellis codes can be found for effective rates between 3.0 and 3.75 bit/T and that are 90° or 180° transparent. The results from a systematic code search using these signal sets are presented.

*This work was supported by NASA Grant NAG5-557.

TRELLIS CODING USING MULTI-DIMENSIONAL QAM SIGNAL SETS

by

Steven S. Pietrobon and Daniel J. Costello, Jr.

June 1989

Submitted to the

1990 IEEE International Symposium on Information Theory

Summary

A systematic method of finding good trellis codes using multi-dimensional QAM signal sets is presented. An important part of these types of trellis codes is in the construction of the multi-dimensional signal sets.

The method used is very similar to that in [1] in which multi-dimensional MPSK signal sets were constructed. That is, we start with a 2-D signal set with $\mathbf{M} = 2^I$ points and form a partition chain such that the minimum squared subset distance (MSSD or δ_i^2) at partition level i is as large as possible. The partition starts at partition level 0 with the whole signal set, dividing each set in two until we are left with M subsets of one point each at partition level I. With rectangular signal sets, it is easily shown that $\delta_{i+1}^2 = 2\delta_i^2$ for $1 \leq i \leq I-2$ and $\delta_I^2 = \infty$.

The next step in forming multi-dimensional signal sets is to take the cartesian product of L of these 2-D signal sets to form a 2L-dimensional (2L-D) signal set and find a partitioning. This is achieved by the use of coset generators which are found from the partitioning of binary block codes. If the 2-D signal set is naturally mapped, a multi-D signal set mapping can be found which has at most I bits affected by a phase rotation out of the total of IL bits used to map the multi-D signal set.

A 16QAM signal set is presented which has these properties. It is shown that only the two lsb's are affected by a 90° phase rotation, while the two msb's are unaffected by a phase rotation. This signal set is then used to construct 4-D, 6-D, and 8-D QAM signal sets which have only 2 bits affected a phase rotation out of the 4L bits used to map the signal set.

Since the multi-D signal sets have only 2 bits affected by a 90° phase rotation (due to the way they are constructed) many of the trellis codes that are found are rotationally invariant to 90° phase rotations.

[1] S. S. Pietrobon, R. H. Deng, A. Lafanechére, G. Ungerboeck, and D. J. Costello, Jr., "Trellis coded multi-dimensional phase modulation", *IEEE Trans. Inform. Theory*, to appear.

Appendix E

Erasurefree Sequential Decoding and Its Application to Trellis Codes

Erasurefree Sequential Decoding and Its Application to Trellis Codes*

Fu-Quan Wang
Daniel J. Costello, Jr.
Dept. of Elec. and Comput. Engr.
University of Notre Dame
Notre Dame, Indiana 46556

presented at
1990 International Symposium
on Information Theory
San Diego, California

January, 1990

*This work was supported by NSF Grant NCR 89-03429 and NASA grant NAG 5-557.

OUTLINE OF PAPER

- Erasurefree Sequential Decoding Algorithms
- Applications to Convolutional Codes
- Sequential Decoding of Trellis Codes
- Erasurefree Decoding of Trellis Codes
- Performance Results
- Conclusions

Why Sequential Decoding ?

- The Viterbi Algorithm (VA) is practical for decoding convolutional codes with small constraint lengths ν.

- Performance (free distance d_{free}) is limited due to small ν.

- Sequential Decoding (SD) can be used with any value of ν.

- Better performance (d_{free}) can be achieved with larger ν.

Problems with Sequential Decoding

- SD's computational effort is a random variable.

- Therefore, some information may be lost due to overflow of the decoder input buffer.

- This results in an erasure probability for SD typically on the order of 10^{-2} to 10^{-3} (Layland and Lushbaugh).

- Complete (erasurefree) decoding may be impossible if a feedback channel is not available.

Goal of This Research

- Propose erasurefree SD algorithms which perform better than the VA and have lower computational requirements.

- Investigate the application of SD to Trellis Codes.

- Some results using conventional SD algorithms with Trellis Codes have been reported by Pottie and Taylor.

Conventional Sequential Decoding Algorithms

- The Fano Algorithm (FA) requires little storage.

- The Stack Algorithm (SA) decodes faster at higher code rates.

- The M- Algorithm (MA) achieves the performance of the VA for asymptotically large SNR.

- The FA requires the least complexity cost to achieve the same performance (for a BER around 10^{-5}). (Anderson and Mohan)

- The FA is prefered in most practical implementations.

Erasurefree Sequential Decoding

• Assume that the information sequence is divided into frames of length L. each terminated by a string of ν zeroes.

• Erasurefree algorithms require that a computational limit C_{lim} be specified for each frame such that:

(1). If the number of computations $C \leq C_{lim}$, a conventional sequential decoding algorithm is used.

(2). If C $> C_{lim}$, a suboptimal decoding algorithm which guarantees complete decoding of the frame is used.

Examples

(1).The Multiple Stack Algorithm (MSA. Chevillat and Costello)-

• Uses one large stack and several smaller stacks.

• Once the main stack is filled, the T best paths are transfered to a secondary stack.

• Once a secondary stack is formed, the decoder can never back up beyond the initial nodes in that stack.

• Additional secondary stacks are formed as needed.

(2). The Erasurefree Fano Algorithm (EFA. new)-

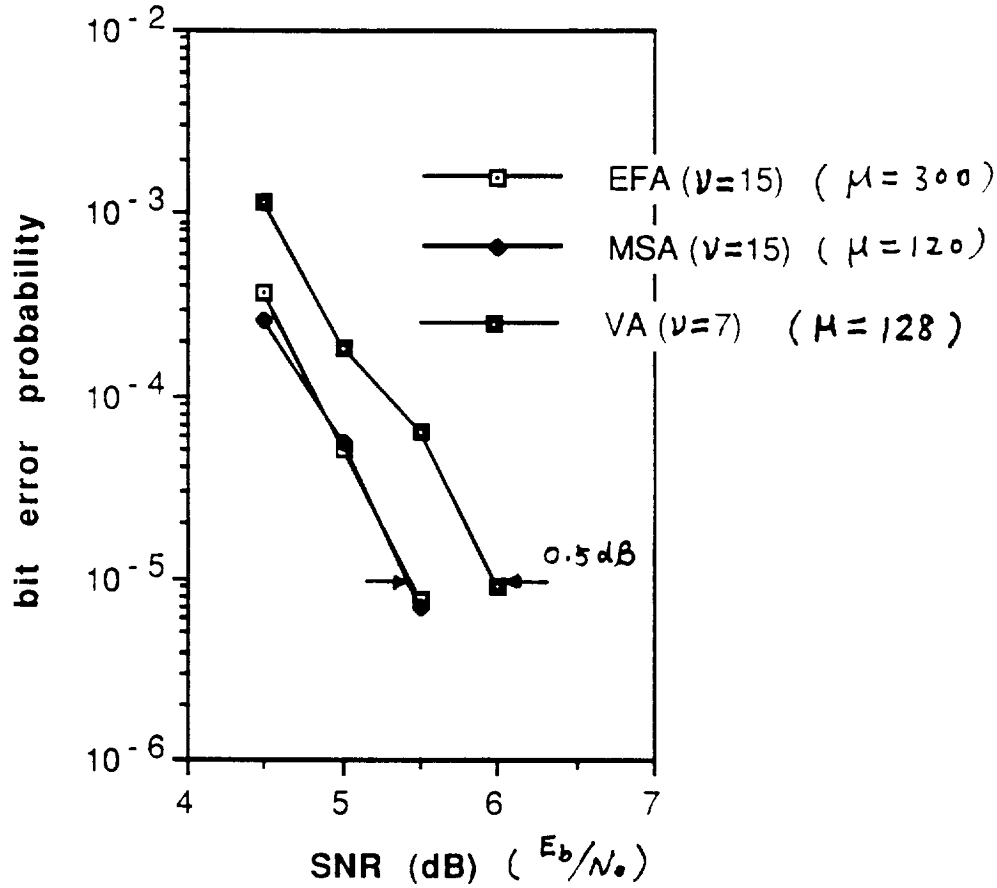
• A predetermined computational limit is set.

• Once this limit is reached. the decoder jumps to the deepest node it has examined thus far (the deepest node must always be stored).

• Decoding resumes at this node and can never back up beyond this node.

• This process is repeated as many times as needed. but each with a smaller computational limit.

Performance Comparison of the MSA, EFA, and VA

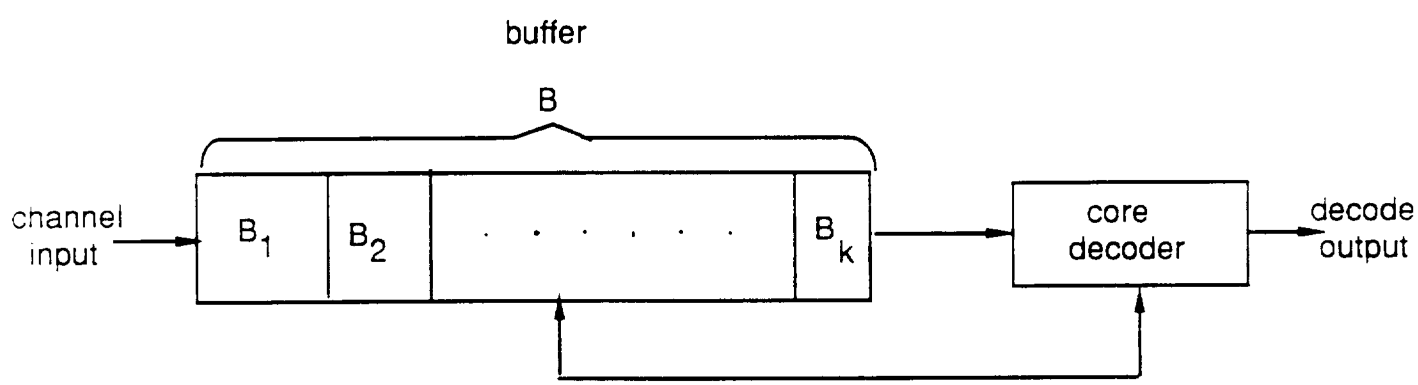


Problems with the MSA and EFA

- Although it is bounded, the number of computations is still a random variable.
- The maximum number of computations per frame, C_{max}, must be large if good performance is desired.
- In order to guarantee erasurefree decoding with a finite buffer, a large speed factor $\mu = C_{max}/(L+\nu)$ is required (say, $\mu \geq 60$ for MSA or $\mu \geq 150$ for EFA).

The Buffer Looking Algorithm (BLA)

Diagram of the BLA



• The input buffer of the decoder is divided into K sections.

• $C_{lim}(j)$ is a computational limit corresponding to the j-th section of the buffer.

• The decoder continuously monitors the buffer state j (number of occupied sections in the buffer).

• If C $\leq C_{lim}(j)$, the BLA works exactly like the FA.

• If C $> C_{lim}(j)$, the BLA works exactly like the EFA.

• If all buffer sections are occupied, the decoder changes parameters (bias) to guarantee the frame is decoded before the buffer overflows.

Erasurefree Decoding Conditions for the BLA

- Let B be the size of the buffer
- Let B_K be the size of the last section of the buffer.
- Let μ be the speed factor of the decoder.
- $B > L + \nu$.
- $B_K > (L + \nu)/\mu$.
- $C_{lim}(K) < (\mu - 1)(L + \nu)$.

Influence of Parameters on Performance

- Number of buffer sections:

Fewer sections allow larger computational limits in each section. (2 is best).

- Buffer size:

Larger buffer size allows more frames to be decoded optimumly.

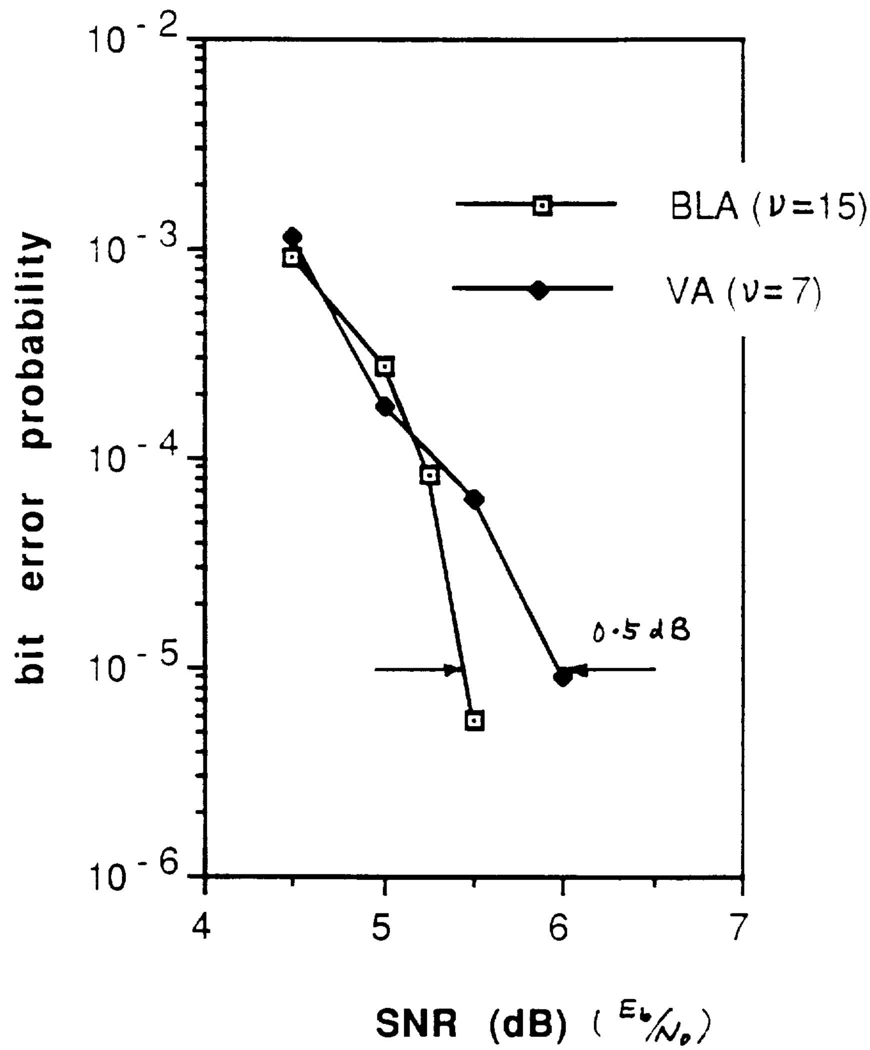
- Speed factor:

Larger speed factor implies more computations are available.

- Frame length:

More data may be decoded suboptimumly for long frames.

Performance Comparison with the VA



BLA: Buffer size = 64 K bits
Speed factor = 4
Frame Length = 256 bits
# of Buffer Sections = 2

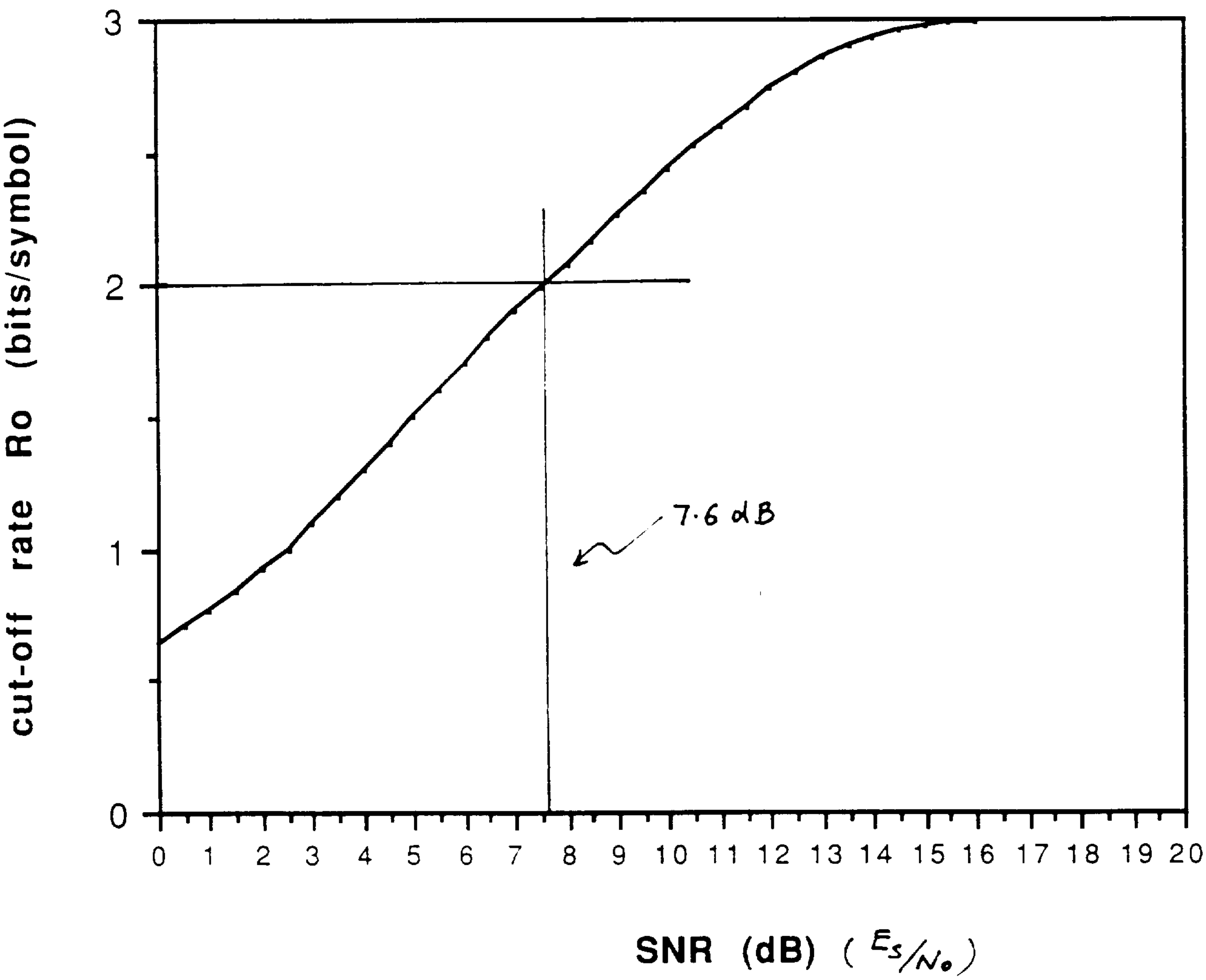
VA: Speed factor = 128

Sequential Decoding of Trellis Codes

- Cut-off rate for two-dimensional signal constellations:

$$R_0 = 2\log_2 K - \log_2\left\{\Sigma_{i=0}^{K-1}\Sigma_{j=0}^{K-1} \exp\left[-\frac{(a_x{}^i - a_x{}^j)^2 + (a_y{}^i - a_y{}^j)^2}{8\sigma^2}\right]\right\}$$

- For 8-PSK, R_0=2 bits/symbol when SNR=7.6 dB.



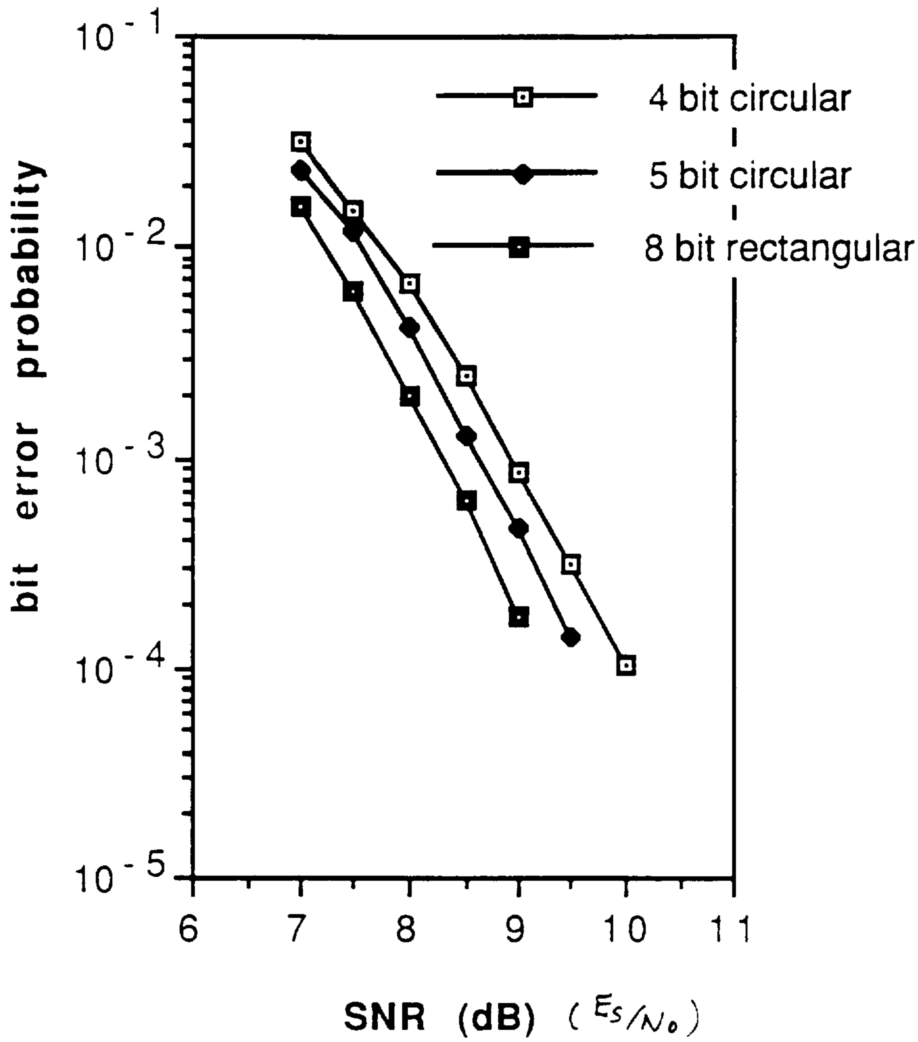
Metric and Threshold Increment for Trellis Codes

- Branch Fano metric:

$$\mathrm{L}(\mathrm{m},\mathrm{y}_i) = \log_2 \frac{\exp(-||y_i - x_{mi}||^2/2\sigma^2)}{\Sigma_{k=1}^{8} p(x_{ki}) \exp(-||y_i - x_{ki}||^2/2\sigma^2)} \\ -3\mathrm{R}$$

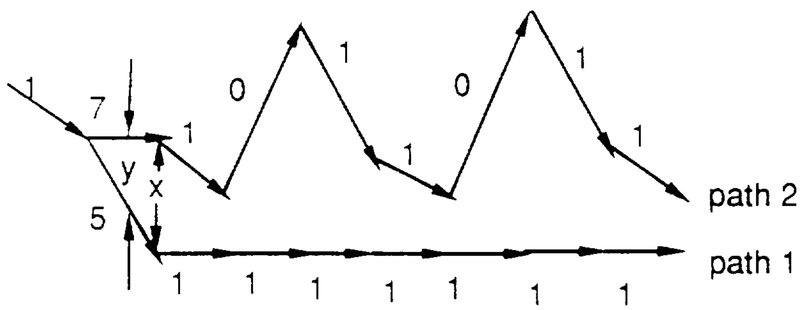
- Unscaled threshold increment Δ should be chosen between 3 to 5 for trellis codes (determined by experiment).

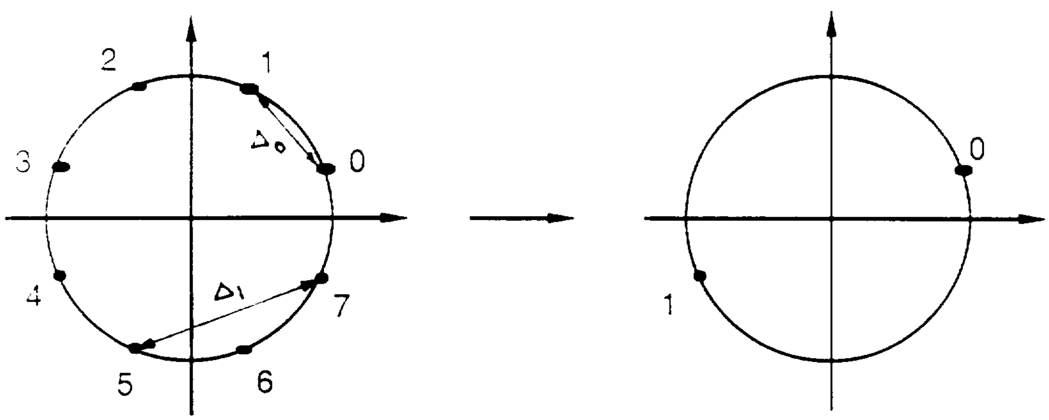
Quantization Schemes



• More than 5 bit circular and 8 bit rectangular quantizations are virtually equivalent to 5 bit circular and 8 bit rectangular respectively.

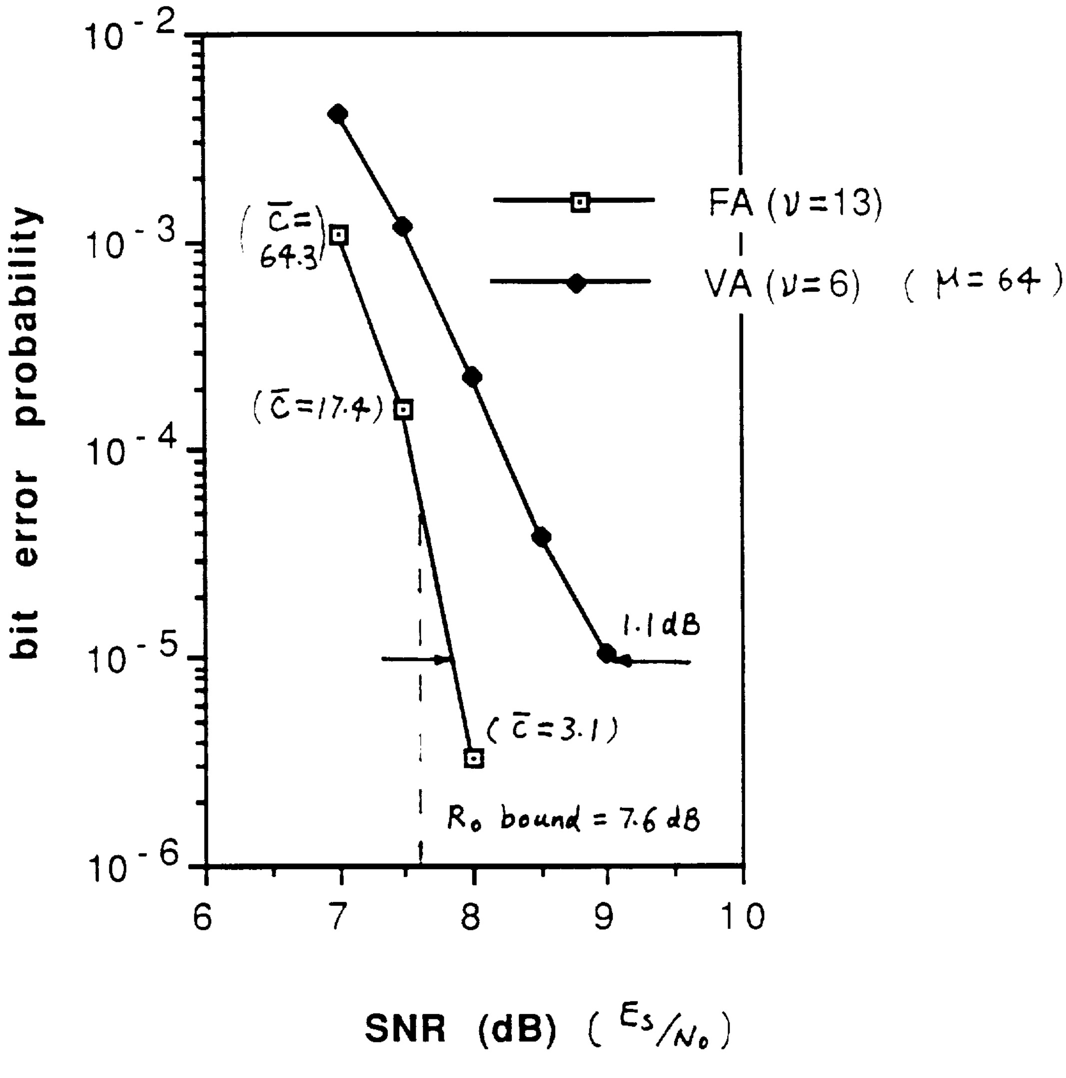
Tail Mapping Must Be Changed for Frame-type Decoding of Trellis Codes





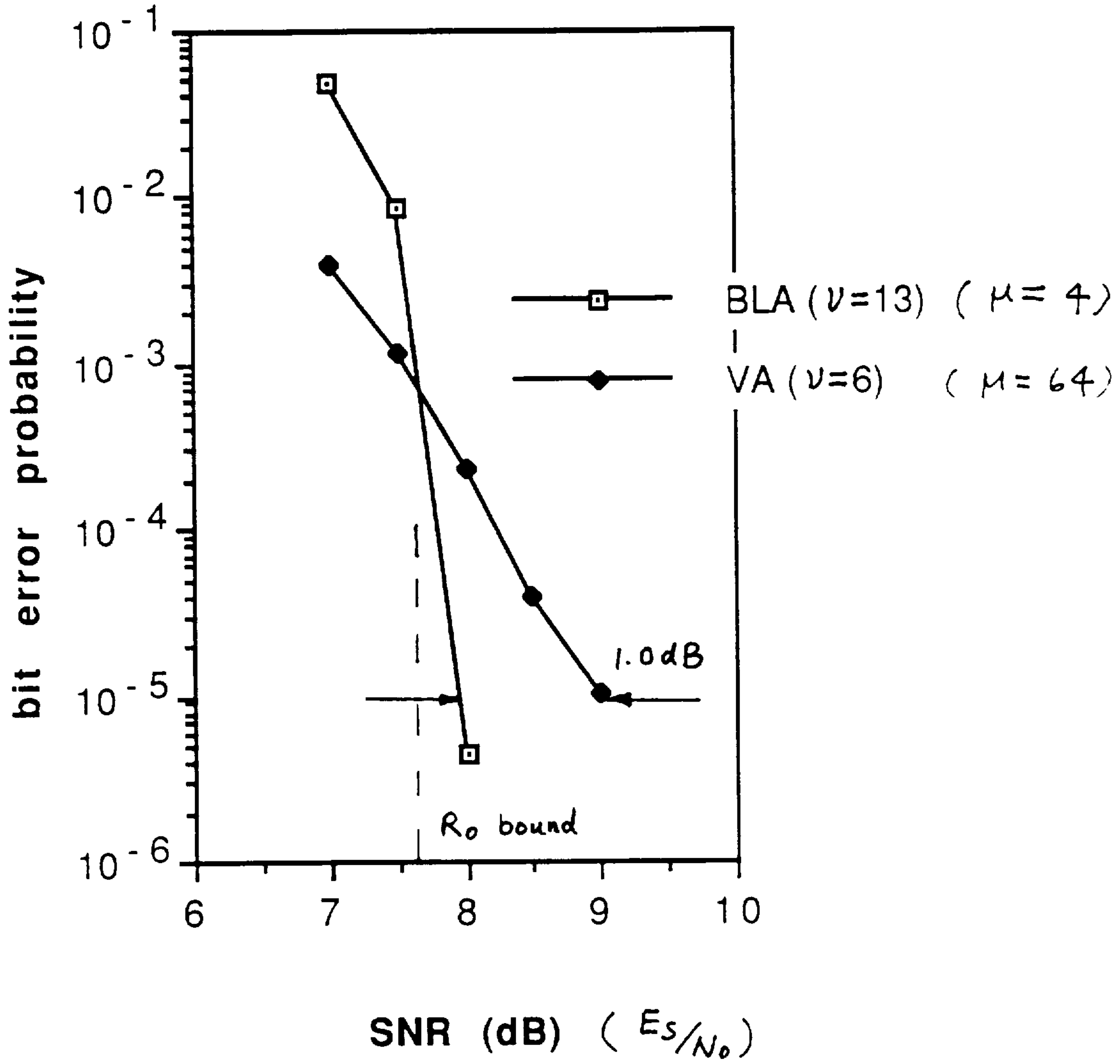
- Ungerboeck 8 - state ($\nu = 3$) code.
- Tail begins at point X and we assume no noise occurs after X.
- Path 1 is the correct path.
- Branch y is corrupted by noise, which makes the decoder follow path 2 (an incorrect path).
- The noise level: $\sqrt{{\Delta_0}^2 + {\Delta_1}^2}/2(= 0.8) < |n| < d_{free}/2(= 1.1)$.
- Natural mapping cannot correct the error in a one constraint length tail.
- This kind of error will dominate in many cases.
- Only 0 and 1 are possible signals in the tail (00X).
- Change the mapping in the tail to achieve a larger distance between signals 0 and 1.

Conventional FA Decoding of Trellis Codes



$\overline{C}$: Average # of Computations

BLA Decoding of Trellis Codes



Conclusions

- Erasurefree sequential decoding algorithms can perform better than the VA with less computational effort.

- SD can work for trellis codes as well as convolutional codes.

- More than 1 dB gain over the VA can be achieved at a BER of 10^{-5} when the BLA is applied to trellis codes (Porath and Aulin code).

Acknowledgement

The authors are grateful to Joseph Nowack and Lance Perez, who provided the simulation results for Viterbi decoding of convolutional and trellis codes, respectively. The first author is indebted to professor Changnian Cai of Beijing University of Posts and Telecommunications for his support and encouragement at the beginning stage of this work.